

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 5, 2004, 07:05:52 ; Search time 368.125 Seconds
(without alignments)
10135.683 Million cell updates/sec

Title: us-09-006-627-3

Perfect score: 736

Sequence: 1 ATTACGTTTGTGTCATCGTA.....TTAAATTTGTATTTTCTA 736

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 3340653 seqs, 2534783454 residues

Total number of hits satisfying chosen parameters: 6681306

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:**

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
18: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
19: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	726	98.6	1440	8	US-08-781-986A-725
C 2	726	98.6	1440	13	Sequence 725, App
C 3	224.8	30.5	495269	16	Sequence 725, App
C 4	224.8	30.5	3011208	16	Sequence 8, Appli
C 5	193	26.2	1908	9	Sequence 2058, App
C 6	137.6	18.7	301208	16	Sequence 27, Appl
C 7	131.8	17.9	690	13	Sequence 2058, App
C 8	130.2	17.7	760	13	Sequence 17336, A
C 9	130.2	17.7	760	13	Sequence 822, App
C 10	130	17.7	31517	9	Sequence 822, App
C 11	128	17.4	1037	16	Sequence 180, App
C 12	126.4	17.2	717	16	Sequence 3163, App
C 13	125.8	17.1	12022	9	Sequence 886, App
C 14	125.6	17.1	699	9	Sequence 117, App
					Sequence 45, Appl

C 15	125.6	17.1	699	15	US-10-084-205-45
C 16	125.6	17.1	699	17	US-10-712-713-45
C 17	125.6	17.1	708	9	US-09-815-242-8902
C 18	125.6	17.1	708	13	US-10-282-122A-8093
C 19	125.6	17.1	9425	8	US-08-781-986A-87
C 20	125.6	17.1	9425	13	US-10-329-624-87
C 21	125.6	17.0	721	13	US-10-282-122A-34637
C 22	124.6	16.9	705	13	US-10-282-122A-10198
C 23	124.4	16.9	696	13	US-10-282-122A-16259
C 24	124	16.8	699	9	US-09-815-242-4621
C 25	124	16.8	708	9	US-09-815-242-8576
C 26	122.8	16.7	319630	16	US-10-398-221-7
C 27	122.2	16.6	678	9	US-09-815-242-7381
C 28	122.2	16.6	678	13	US-10-335-977-2707
C 29	122.2	16.6	678	13	US-10-335-977-2708
C 30	120	16.3	663	16	US-10-398-221-732
C 31	120	16.3	663	16	US-10-398-221-2755
C 32	119.6	16.2	3636	13	US-10-158-844-78
C 33	119	16.2	678	10	US-09-882-227-485
C 34	118.8	16.1	702	13	US-10-282-122A-6376
C 35	118.8	16.1	705	9	US-09-815-242-6651
C 36	118.8	16.1	7294	9	US-09-070-927A-63
C 37	116.8	15.9	699	9	US-09-815-242-3858
C 38	116.4	15.8	2400	10	US-09-877-963-1
C 39	116.4	15.8	2400	10	US-09-877-963-5
C 40	115.6	15.7	699	9	US-09-974-300-4345
C 41	113.4	15.4	785	13	US-10-282-122A-21390
C 42	113.2	15.4	699	13	US-10-282-122A-17079
C 43	110.2	15.0	1830121	15	US-10-329-960-1
C 44	110.2	15.0	1830121	16	US-10-329-960-1
C 45	109.8	14.9	720	15	US-10-287-274-196

ALIGNMENTS

RESULT 1

US-08-781-986A-725/c
Sequence 725, Application US/08/781/986A
Publication No. US20030054436A1
GENERAL INFORMATION:
APPLICANT: Charles Kunsch
TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
NUMBER OF SEQUENCES: 5255
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/781,986A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Benson, Bob
REGISTRATION NUMBER: 30,446
REFERENCE/DOCKET NUMBER: PB248PP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 725:
SEQUENCE CHARACTERISTICS:
LENGTH: 1440 base pairs

Sequence 45, Appl
Sequence 85, Appl
Sequence 4502, Ap
Sequence 8093, Ap
Sequence 87, Appl
Sequence 87, Appl
Sequence 34637, A
Sequence 10198, A
Sequence 16259, A
Sequence 4621, Ap
Sequence 8576, Ap
Sequence 7, Appl
Sequence 2707, Ap
Sequence 2708, Ap
Sequence 732, App
Sequence 732, App
Sequence 7255, Ap
Sequence 78, Appl
Sequence 485, App
Sequence 6376, Ap
Sequence 6651, Ap
Sequence 63, Appl
Sequence 3858, Ap
Sequence 1, Appl
Sequence 5, Appl
Sequence 4345, Ap
Sequence 21390, A
Sequence 17079, A
Sequence 1, Appl
Sequence 1, Appl
Sequence 196, App

TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-781-986A-725

Query Match 98.6%; Score 726; DB 8; Length 1440;
Best Local Similarity 99.3%; Pred. No. 8.8e-141;
Matches 726; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

```
Qy 1 ATTACGTTTGTGTCATCGTATCATACCAAGCCACGAACTGTTTCAATCAATTTGTC 60
Db |||
Qy 734 ATTACGTTTGTGTCATCGTATCATACCAAGCCACGAACTGTTTCAATCAATTTGTC 675
Db |||
Qy 61 ACGATCGTATGTTTAACTGTTTTCGTAATATCTTATTAACATCTAGCAATTTGT 120
Db |||
Qy 674 ACGATCGTATGTTTAACTGTTTTCGTAATATCTTATTAACATCTAGCAATTTGT 615
Db |||
Qy 121 TTCTACTTTCACTATTATAAACCACATACATGATTTTAAATTTGTTCCCGTTGCATAACATG 180
Db |||
Qy 614 TTCTACTTTCACTATTATAAACCACATACATGATTTTAAATTTGTTCCCGTTGCATAACATG 555
Db |||
Qy 181 GTTTTATTTTCACTAGTAGAGATATAGTAAATCATCTCTGTTTGTAAATTTCAATTTTC 240
Db |||
Qy 554 GTTTTATTTTCACTAGTAGAGATATAGTAAATCATCTCTGTTTGTAAATTTCAATTTTC 495
Db |||
Qy 241 TCGCCATTTTACGCTGCTTAAAGCGTTTCTTATCAATTTGTAATACCGTTGACATCGAT 300
Db |||
Qy 494 TCGCCATTTTACGCTGCTTAAAGCGTTTCTTATCAATTTGTAATACCGTTGACATCGAT 435
Db |||
Qy 301 AATATCTTTTGTGGCTGACGACGATAAATTTGCAAGAACTTTGCTAAAAAGTTCTTCAAT 360
Db |||
Qy 434 AATATCTTTTGTGGCTGACGACGATAAATTTGCAAGAACTTTGCTAAAAAGTTCTTCAAT 375
Db |||
Qy 361 ATCAAAACGGCTTAACTATATAATCGTCTGCAACGTAATCAAGCCACGAACTTTGTCATA 420
Db |||
Qy 374 ATCAAAACGGCTTAACTATATAATCGTCTGCAACGTAATCAAGCCACGAACTTTGTCATA 315
Db |||
Qy 421 CGTATCACTTTTTCGCTGTAATTAATGATAGGTGATGTTGTTGTTGTTGTTGCG 480
Db |||
Qy 314 CGTATCACTTTTTCGCTGTAATTAATGATAGGTGATGTTGTTGTTGTTGTTGCG 255
Db |||
Qy 481 ACAAAATTTCTAAGCCATTAATGACGCAACATTAATCTAATATGATTAATCAATAGTA 540
Db |||
Qy 254 ACAAAATTTCTAAGCCATTAATGACGCAACATTAATCTAATATGATTAATCAATAGTA 195
Db |||
Qy 541 ATGCTAAGCGCTTTATCAAAACGCTCTGTCCATCATCTCTGTGTCCATGTAATTT 600
Db |||
Qy 194 ATGCTAAGCGCTTTATCAAAACGCTCTGTCCATCATCTCTGTGTCCATGTAATTT 135
Db |||
Qy 601 TTCAATGTTGAGTTTCCAAATTCAGAAATCTTGTGTAAGTTTGTCTTCTACTATTAA 660
Db |||
Qy 134 TTCAATGTTGAGTTTCCAAATTCAGAAATCTTGTGTAAGTTTGTCTTCTACTATTAA 75
Db |||
Qy 661 AATTTGGCTCAATTTGTACCTCATATACGACTTTTCTAATAAGGTAATATATTAA 720
Db |||
Qy 74 AATTTGGCTCAATTTGTACCTCATATACGACTTTTCTAATAAGGTAATATATTAA 15
Db |||
Qy 721 AATTTGTATTT 731
Db |||
Qy 14 ATANTGNAATT 4
```

RESULT 2
US-10-329-624-725/c
; Sequence 725, Application US/10329624
; Publication No. US20040043037A1
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; Gil H. Choi
; Patrick S. Dillon
; Craig A. Rosen
; Steven C. Barash
; Michael R. Fannon

TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
NUMBER OF SEQUENCES: 5256
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICANT NUMBER: US/10/329,624
FILING DATE: 27-Dec-2002
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/956,171
FILING DATE: October 20, 1997
APPLICATION NUMBER: 60/009,861
FILING DATE: January 5, 1996
APPLICATION NUMBER: 08/781,986
FILING DATE: January 3, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Mark J. Hyman
REGISTRATION NUMBER: 46,789
REFERENCE/DOCKET NUMBER: PB248P1D1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (240) 314-1224
TELEFAX: (301) 309-8439
INFORMATION FOR SEQ ID NO: 725:
SEQUENCE CHARACTERISTICS:
LENGTH: 1440 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 725:
US-10-329-624-725

Query Match 98.6%; Score 726; DB 13; Length 1440;
Best Local Similarity 99.3%; Pred. No. 8.8e-141;
Matches 726; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

```
Qy 1 ATTACGTTTGTGTCATCGTATCATACCAAGCCACGAACTGTTTCAATCAATTTGTC 60
Db |||
Qy 734 ATTACGTTTGTGTCATCGTATCATACCAAGCCACGAACTGTTTCAATCAATTTGTC 675
Db |||
Qy 61 ACGATCGTATGTTTAACTGTTTTCGTAATATCTTATTAACATCTAGCAATTTGT 120
Db |||
Qy 674 ACGATCGTATGTTTAACTGTTTTCGTAATATCTTATTAACATCTAGCAATTTGT 615
Db |||
Qy 121 TTCTACTTTCACTATTATAAACCACATACATGATTTTAAATTTGTTCCCGTTGCATAACATG 180
Db |||
Qy 614 TTCTACTTTCACTATTATAAACCACATACATGATTTTAAATTTGTTCCCGTTGCATAACATG 555
Db |||
Qy 181 GTTTTATTTTCACTAGTAGAGATATAGTAAATCATCTCTGTTTGTAAATTTCAATTTTC 240
Db |||
Qy 554 GTTTTATTTTCACTAGTAGAGATATAGTAAATCATCTCTGTTTGTAAATTTCAATTTTC 495
Db |||
Qy 241 TCGCCATTTTACGCTGCTTAAAGCGTTTCTTATCAATTTGTAATACCGTTGACATCGAT 300
Db |||
Qy 494 TCGCCATTTTACGCTGCTTAAAGCGTTTCTTATCAATTTGTAATACCGTTGACATCGAT 435
Db |||
Qy 301 AATATCTTTTGTGGCTGACGACGATAAATTTGCAAGAACTTTGCTAAAAAGTTCTTCAAT 360
Db |||
Qy 434 AATATCTTTTGTGGCTGACGACGATAAATTTGCAAGAACTTTGCTAAAAAGTTCTTCAAT 375
Db |||
Qy 361 ATCAAAACGGCTTAACTATATAATCGTCTGCAACGTAATCAAGCCACGAACTTTGTCATA 420
Db |||
Qy 374 ATCAAAACGGCTTAACTATATAATCGTCTGCAACGTAATCAAGCCACGAACTTTGTCATA 315
Db |||
Qy 421 CGTATCACTTTTTCGCTGTAATTAATGATAGGTGATGTTGTTGTTGTTGTTGCG 480
```



```
Db 314 CGTATCACTTTTCGCTGTAATTTATTAATCATAGGTGTAGATTGTGTGTTAAATTTGCG 255
Qy 481 ACAAAATTTCTAAGCCATTAAATGACGCAACATTAAATCTAATATGATTAATCATAGTA 540
Db 254 ACAAAATTTCTAAGCCATTAAATGACGCAACATTAAATCTAATATGATTAATCATAGTA 195
Qy 541 ATGGCTAAGCGCTTTATCTAAACCGTCTTGTCATCATCTGTCGTCACCATTTGTAAT 600
Db 194 ATGGCTAAGCGCTTTATCTAAACCGTCTTGTCATCATCTGTCGTCACCATTTGTAAT 135
Qy 601 TTCATGTGTGAGTTCCTAATCAAGAAATCTTGCTAAGTTTGTTCATCTTCTCATATTA 660
Db 134 TTCATGTGTGAGTTCCTAATCAAGAAATCTTGCTAAGTTTGTTCATCTTCTCATATTA 75
Qy 661 AATTTGCGTCATTTGTACACCTCATATTACGACTTTTCTAATAGGTAATATTTTAA 720
Db 74 AATTTGCGTCATTTGTACACCTCATATTACGACTTTTCTAATAGGTAATATTTTAA 15
Qy 721 AATTTGTAATTT 731
Db 14 ATANTGNATTT 4

RESULT 3
US-10-398-221-8
; Sequence 8, Application US/10398221
; Publication No. US20040018514A1
; GENERAL INFORMATION:
; APPLICANT: KUNST, Frederik
; APPLICANT: GLASER, Philippe
; TITLE OF INVENTION: Listeria innocua, genome and applications
; FILE REFERENCE: 344 702 - US
; CURRENT APPLICATION NUMBER: US/10/398, 221
; CURRENT FILING DATE: 2003-03-27
; PRIOR APPLICATION NUMBER: PCT/FR 01/03 061
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: FR 00/12 697
; PRIOR FILING DATE: 2000-10-04
; NUMBER OF SEQ ID NOS: 4025
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 8
; LENGTH: 495269
; TYPE: DNA
; ORGANISM: Listeria innocua
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(end)
; OTHER INFORMATION: n can be any nucleotide: a, g, c or t/u
US-10-398-221-8

Query Match 30.5%; Score 224.8; DB 16; Length 495269;
Best Local Similarity 59.4%; Pred. No. 7.7e-36; Indels 18; Gaps 1;
Matches 410; Conservative 0; Mismatches 262;

Qy 4 TACGTTTTGTTCATCGTATCACATACCCAGCCAGCACTGTTTCAATCATTTTGTGCAG 63
Db 348840 TAGTTGTATGTACGCATCACATACCCAGTCCCGCACTGTTTGGATGTAACTTTCTTC 348899
Qy 64 ATCGTATGGTTTTAACTGTTTGTAAATATCTTATATAAATCATCTACGACATTTGTTTC 123
Db 348900 GTCAGGATGATCAATTTTATTTTCGCAAGTAACGAACATATACATCTACTACATTTGTTTC 348959
Qy 124 TACTTCACTATTAAACCCCATCATGATTTAAATTTGTTCCGTTTGCATAACATGGTT 183
Db 348960 TACTTCTGTTTCAATACCCCATCATCTTGTAAAGTAATACITCCCGAGTAAGAACATTT 349019
Qy 184 TTTATTTTCAGCTAGAAGATATAGTAATCATCTCTGTTTTTGTAAATTCATTTCTGC 243
Db 349020 GACATTTTCCATTAAGTAAGTAGTGTTCATCTCGGTTTGTGTTAAATCAATATTTTC 349079
Qy 244 GCCATTTACCGTCACTTTTAAAGCGTTTCTTATCAATTTGTAATACCGTTGACATCGAAT 303
Db 14 ATANTGNATTT 4
```

```
Db 349080 TTCAATCGGTTTTCAAAATCCGATTTTCTTTTCAACGATTAAGTTGCGGTATTGTAGCGT 349139
Qy 304 -----ATCCTTTTGTGGCTGACGACGTAAATTTGCAGAAATTTCTTGC 345
Db 349140 TGTTTGTTTAGCAGATTGTTCTGCAATTTTCTACTCGGGTAATAGTGAACGCAACGTCG 349199
Qy 346 TAAAGTTTCTCAATATCAAAACGGCTTAACATATATAATCGTCTGCACCGTAATCAAGCCC 405
Db 349200 AAGTAGTTCTTCAATTCGAAATGGCTTAAACGATATAATCATCTGCTCCGTTGGTCAAGTCC 349259
Qy 406 AGCAACTTTGTCTATCATCTATCACTTTTTCGCTGTAATTTATTAATGATAGGTGTAGATTGTTG 465
Db 349260 AGATACCGGATCAATCACAGAGTCTCGCGCAGTTTATCATTAATGATGGGTGTTGTTTCAAC 349319
Qy 466 TTGTCTAATTTTGGCAAAATTTCTAAGCCATTAATTTGACGCCAACATTAATCTAATAT 525
Db 349320 TTGGCGTACAGGACGACAAACCTCTACCCGTTTAAATGTTGTAACATTAAGTCTAGTAA 349379
Qy 526 GATTAATATCATAGTAAGGCTTAAGCGCTTTTATCTAATAACCGTCTTGTCCATCATCTCTGT 585
Db 349380 AATAGCATCCCATTTCTTCAATTTAGTGGAGTTCTAGCCAGCAGTCCATCATTTAGCAAC 349439
Qy 586 GTCCACATTTGTAATTTTCTATGTGTGAGTTCCAAATTCAGGAATCTTCTAGTAAATTTGTTTC 645
Db 349440 AGCAGTTTCAATAATTTTCGTTGTTGTAATTTCTAGTTCAATAAAGCGTGTAGTTTTC 349499
Qy 646 ATCTTCTACTATTAAAAATTTTGGCGTCATTTG 675
Db 349500 ATCTTCTACGATTAATTTCTATTTCAATTTG 349529

RESULT 4
US-10-398-221-2058/c
; Sequence 2058, Application US/10398221
; Publication No. US20040018514A1
; GENERAL INFORMATION:
; APPLICANT: KUNST, Frederik
; APPLICANT: GLASER, Philippe
; TITLE OF INVENTION: Listeria innocua, genome and applications
; FILE REFERENCE: 344 702 - US
; CURRENT APPLICATION NUMBER: US/10/398, 221
; CURRENT FILING DATE: 2003-03-27
; PRIOR APPLICATION NUMBER: PCT/FR 01/03 061
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: FR 00/12 697
; PRIOR FILING DATE: 2000-10-04
; NUMBER OF SEQ ID NOS: 4025
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2058
; LENGTH: 3011208
; TYPE: DNA
; ORGANISM: Listeria innocua
US-10-398-221-2058

Query Match 30.5%; Score 224.8; DB 16; Length 3011208;
Best Local Similarity 59.4%; Pred. No. 1.5e-35;
Matches 410; Conservative 0; Mismatches 262; Indels 18; Gaps 1;

Qy 4 TACGTTTTGTTCATCGTATCACATACCCAGCCAGCACTGTTTCAATCATTTTGTGCAG 63
Db 1410111 TAGTTGTATGTACGCATCACATACCCAGTCCCGCACTGTTTGGATGTAACTTTCTTC 1410052
Qy 64 ATCGTATGGTTTTAACTGTTTGTAAATATCTTATATAAATCATCTACGACATTTGTTTC 123
Db 1410051 GTCAGGATGATCAATTTTATTTTCGCAAGTAACGAACATATACATCTACTACATTTGTTTC 1409992
Qy 124 TACTTCACTATTAAACCCCATCATGATTTAAATTTGTTCCGTTTGCATAACATGGTT 183
Db 1409991 TACTTCTGTTTCAATACCCCATCATCTTGTAAAGTAATACITCCCGAGTAAGAACATTT 1409932
Qy 184 TTTATTTTCAGCTAGAAGATATAGTAATCATCTCTGTTTTTGTAAATTCATTTCTGC 243
Db 1409931 GACATTTTCCATTAAGTAAGTAGTGTTCATCTCGCGTTTGTGTTAAATCAATAATTTTC 1409872
```


RESULT 7
US-10-282-122A-17336/c
; Sequence 17336, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos

Db 279 TTCAACCTTGTCATTTTCCATATTTTACAGTAAGCAATATTAACAGGAATATTTTAA 220
QY 464 -----TGTTGCTAATTTTGGGCAAAATTTCTAAGCCATTAATTTGACGGCAACATTA 516
Db 219 GTCCTTATCACTTCTTATTTTCTCAAAACTTCTATTCACATTTATATTTGGTAACATCCA 160
QY 517 ATCTAATATGATTAATCATAGTAAGTGAAGCGCTTTATCTAAGCCGCTTGTGCTCATC 576
Db 159 ATCTAAGATTAATTAATCTCGTTTAAATTTCCCTTAGCCCTTTATAAATCCCGTCAATCCATC 100
QY 577 ATACTCTGTGTCACATTTGTAATTTTTCATGTGTGAGTTCCAAATTCAGAAATCTTCTCTAA 636
Db 99 ATATGAATTAACCTTCTGTAATTTTGACACCTCCAAATTTGAATTTTAATAGCTCAACTAT 40
QY 637 GTTTTGTGTCATCTTCTACTATTAATAATTTGCGT 669
Db 39 ATGCATTTTCATGCTCTATATAACAAGCACTTTAGT 7

RESULT 8

US-08-781-986A-822/c
; Sequence 822, Application US/08781986A
; Publication No. US2003005436A1
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5255
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781,986A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Benson, Bob
; REGISTRATION NUMBER: 30,446
; REFERENCE/DOCKET NUMBER: PB248PP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 822:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 760 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-781-986A-822

Query Match 17.7%; Score 130.2; DB 8; Length 760;
Best Local Similarity 52.6%; Pred.No. 2.7e-17;
Matches 358; Conservative 0; Mismatches 288; Indels 35; Gaps 2;
QY 25 ATACCAACCCAGCACTGTTCAATCATTTGTGACGATC-----GTATGGTTTAAAC 79
Db 735 ATACCAACCCAGCACTGTTCAATCATATGCGAGCTTCACTAGACACAGATTTAA 676
QY 80 TTGTTTCGTAATATCTTATATAACATCAATGACATCTGTTTCTACTTCACTATTATAA 139
Db 675 CTTTCTCTAAGCGTTTAAACATGATATCAACTGTTCTTAAATCACCATAGAAATTCATA 616

QY 140 CCCCATACATGATTTAAATTTGTTCCCGTTCGATAACAATGTTTATTTTATTTTTCAGCTAGA 199
Db 615 TGCCAAACCTCTTTTAAATTAATTTGTTACCGTCAAAATACCTTTATTTGTTGTTTATAGCTAA 556
QY 200 AGATATAGTAATATCATCTCTGTTTGTGTTAATTTCAATTTCTGCGCAATTTACCGTCACT 259
Db 555 TATATTAATAATTCGTACTCTTTAGGAGTCAAAATTAACCTTCTTGATTAATCAGCAAGTACG 496
QY 260 TTAAGCGCTTCTTATCAATTTGTAATACCGTTGACATCGATTAATATCTCTTTTG----- 312
Db 495 CGATGTGCATCATTTATTTCTAAATGTTTAAATTTCAATCACATCACGTGCGTGAGGT 436
QY 313 -----TGCTGACGACGTAAATTTGACGATTTCTTTAGAGTGCTTTAACTCTTAAAGACT 349
Db 435 TCGCTTTGTTTACAGTTGTAGATTTGCGTTCTTTTAGAGTGCTTTAACTCTTAAAGACT 376
QY 350 AGTTCTTCAATATCAAAACGGCTTAACTATATATATATGCTCTGACCGGTAAATCAAGCCGACGA 409
Db 375 ACTTCTCTTGGTGAATAATGTTTGAAGATATAATCATCTGCACAGATTTCAAAACCTTCA 316
QY 410 ACTTTGTATACGTATCACTTTTTCGTTGTAATTAATTAATGATAGGTAGTAGTGTGTTGTT 469
Db 315 ACACGGTTTGTGTTTCTTACCTTTAGCAGTCAACATAATAATCGGTGTTTGTGTTTATGTTCA 256
QY 470 CTAATTTTTCGACAAATTTCTAAGCATTAAATTTGACGCAACATTAATTAATCTAATATGATT 529
Db 255 CGCAATTTAGTTGCGACCTGGATACATCCATTTGAGGCAACATTAATTAATCTAGTAGTAG 196
QY 530 AAATCATAGTAATGCTAAGCGCTTTATCTAAACCGTCTTTGTCATCATCATCTCTGTGTC 589
Db 195 CAAGCATATATTTCTCCATTCGAATTCATAAGCGCTCTTGGCCATTTACTTGTCTCATGG 136
QY 590 ACATTGTAATTTTATGTTGAGTTCCTCAATTCAGAAATCTTCTGCTAAGTTTGTGTTTCACT 649
Db 135 ATTCAAAAGATTTCTCTTTCTAATATACATTTTAAAGTAATCTTCTGATTTCTATCTCATCA 76
QY 650 TCTACTATTAAATTTGCGTC 670
Db 75 TCTACGATAAGTATTTCGTTTC 55

RESULT 9

US-10-329-624-822/c
; Sequence 822, Application US/10329624
; Publication No. US20040043037A1
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; Gil H. Choi
; Patrick S. Dillon
; Craig A. Rosen
; Steven C. Barash
; Michael R. Fannon
; ADDRESS: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/329,624
; FILING DATE: 27-Dec-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/956,171
; FILING DATE: October 20, 1997
; APPLICATION NUMBER: 60/009,861

; FILING DATE: January 5, 1996
; APPLICATION NUMBER: 08/781,986
; FILING DATE: January 3, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mark J. Hyman
; REGISTRATION NUMBER: 46,789
; REFERENCE/DOCKET NUMBER: PB248P1D1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (240) 314-1224
; TELEFAX: (301) 309-8439
; INFORMATION FOR SEQ ID NO: 822:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 760 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 822:
US-10-329-624-822
Query Match 17.7%; Score 130.2; DB 13; Length 760;
Best Local Similarity 52.6%; Pred. No. 2.7e-17;
Matches 358; Conservative 0; Mismatches 288; Indels 35; Gaps 2;
QY 25 ATACCAACGCCACGACTGTTCAATCAATTTGTCAGATC-----GTATGGTTTAAAC 79
DB 735 ATACCAACGCCACGACTGTTGAATCATATGCGCAGCTTCATAGACACACGATTTAA 676
QY 80 TGTGTTGCGAAATATCTATATAAATCACTACGACATTTGTTCTACTTCTACTATTATAA 139
DB 675 CTTTCTCTAACGGTTAACATGATATCACTGTTCTTAATCAATCAATAGAAATTCATAA 616
QY 140 CCCCATACATGATTTAAATTTGTTCCGGTTGCATACATGTTGTTTATTTTTCAGCTAGA 199
DB 615 TGCCAAACTCTTTTAAATTTGTTACGGTCAATATCTTTATTTGGTGTTTTACTTAA 556
QY 200 AGATATAGTAATCATCTGTTGTTTGTAAATCAATTTTCGCCCAATTTACCGTCACT 259
DB 555 TATATTAATAATTCGTACTCTTTAGAGTCAAAATTAATCTTTGATTTATCAGCAAGTAGC 496
QY 260 TTAAAGAGGTTCTTATCAATGTAATACCGTTGACATGATATATCTTTTG----- 312
DB 495 CGATGTGCATCATATCTATTTCTAAATGTTTAAATTCATCAATCAGTCGCGTAGGT 436
QY 313 -----TGCGTCAGCAGCTAAATTTGACGAAATTTGACGAAATTTCTTCTTAA 349
DB 435 TCGCTTTGTTCTACAGTTGTAGATTGCGTTCTTCTTAGAAGTCTTTTAACTCTTAAGACT 376
QY 350 AGTTCTTCAATATCAACGGCTTAATCAATATATCGTCTGACCGCTAATCAAGCCGACGA 409
DB 375 ACTTCTCTTGGTGAATATGGTTGACGATATAATCATCTGCACCAAGATTCAAAACCTTCA 316
QY 410 ACTTTGTATAGTATCACTTTTCGCTGTAATTAATTAATGATAGGTAGATTGTTGT 469
DB 315 ACACGGTTGTTTCTTCACTTTAGCAGTCAACATAATTAATCGGTGTTGTTTATGTCA 256
QY 470 CTAATTTTGCAGCAATTTCTAAGCCATTAATTTGACCGCAACATTTAAATCTAATATGATT 529
DB 255 CGCAATTTAGTGGCACCTGGATACCATCCATTTCCAGCAACATTTAATCTAGTAGTAG 196
QY 530 AAATCATATAGTATGGCTAAGCGCTTATCTAAACCGCTCTTGCCATCATCTCTGTGTCC 589
DB 195 CAAGCATATAATTTCTCCATTGCAAGTTTCATAAGCCTCTTGCCCATTAATCTTGTCTATGG 136
QY 590 ACATGTATTTTCAATGTGATGTTCCAAATTCAGAAATCTTGCTAAGTTTGTTCATCT 649
DB 135 ATTTCAAAAGATTTCTTTCTTAAATACATTTTAAAGTAATCTTCTGATTTCTATCTCATCA 76
QY 650 TCTACTATTTAAATTTGCGTC 670
DB 75 TCTAGGATAAGTATTTTCGTTTC 55

RESULT 10

US-09-070-927A-180
; Sequence 180, Application US/09070927A
; Patent No. US20020120116A1
; GENERAL INFORMATION:
; APPLICANT: Charles A. Kunsch
; Patrick J. Dillon
; Steven Barash
; TITLE OF INVENTION: Enterococcus faecialis Polynucleotides and Polypeptides
; NUMBER OF SEQUENCES: 982
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/070,927A
; FILING DATE: 04-May-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/046,655
; FILING DATE: 1997-05-16
; APPLICATION NUMBER: 60/044,031
; FILING DATE: 1997-05-06
; APPLICATION NUMBER: 60/066,009
; FILING DATE: 1997-11-14
; ATTORNEY/AGENT INFORMATION:
; NAME: Kenley K. Hoover
; REGISTRATION NUMBER: 40,302
; REFERENCE/DOCKET NUMBER: PB369
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 180:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 31517 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 180:
US-09-070-927A-180
Query Match 17.7%; Score 130; DB 9; Length 31517;
Best Local Similarity 53.4%; Pred. No. 1.2e-16;
Matches 311; Conservative 0; Mismatches 250; Indels 21; Gaps 1;
QY 133 ATTATAACCCCATACATGATTTAAATTTGTTCCGTTGCATACATGTTTATTTTATTTTC 192
DB 1332 ATCAAAACCCACACACTTTCAAAAGTTCTCTTTATTTCCACACACGGTTAGGATGTC 1391
QY 193 AGCTAGAGATATAGTAAATCACTCTGTTTTTGTAAATTCATTTCTGCGCCATTTC 252
DB 1392 CATCAAGAAAACACGCAATTTGATTTCTTGTGTTGTAATAATACTTCATTTCTTAAAC 1451
QY 253 CGTCACCTTTAAACGGTTCTTATCAATTTGTAATACCGTTGACATCGATAATATCCTTTTG 312
DB 1452 AAACACTTTGCGCGCAGTGTGACAGCAATTTCAATTTTATTAGTATTTGTTGGTGG 1511
QY 313 TGGCTG-----ACGACGTAATAATTGCAGCAATTTCTTGTCTTAAAG 351
DB 1512 TTGTTCTGCTTGTCTTAACAATTCATAACGGTTTCATATGGGCTTTTGTCTGAGCGACTAA 1571
QY 352 TTCTTCAATATCAACGGCTTAATATATATCGTCTGACCGTAATCAAGCCGCAAC 411
DB 1572 TTCAATGGGACTAAACGGCTTGATTTATATATATCGCAGCTTAAGCAAGCCGCAAC 1631
QY 412 TTTGTCATACGTATCACTTTTTCGCTGTAATAATATAGTGTAGATTGTTGTGCT 471

Db 1632 TTATCAATATCTCTTTTATAGCGAACAATCATGATCGCGTGTGCTTTTTTTTTCG 1691
 Qy 472 AATTTGGGACAAATTTCTAAGCATTAATTTAGCGCAACATTAATCTAATATGATTA 531
 Db 1692 AATTTGCACGACAGATTTTCAAAACCATCATATAGTTGGCAGCATCACATCTAAAAATAATTA 1751
 Qy 532 ATCATAGTAATGGCTAAGCGCTTTATCTAAACCGTCTTTGCCATCATCTCTGTGTCAC 591
 Db 1752 ATCAAGGGCTCATTTAAAGCTGCTTCCATCTCTTTTTCATCATGTTCAATGTGAC 1811
 Qy 592 ATTGTAATTTTCATGTGTAGTTCCCAATTCAGAAATCTTTGCTAAGTTTGTTCATCTTC 651
 Db 1812 AGTCATATCTTAATTTCTTAATAATCTTTTTCGCAATCAGCAATGCTTGGATCATCTTC 1871
 Qy 652 TACTATTAATAATTTGGCTCATTTGTACACCTCATATATACAC 693
 Db 1872 AATAATTAATACATGTGGTTTCATTTTTCATCTCCTCAACAAC 1913

RESULT 11

US-10-398-221-3163
 ; Sequence 3163, Application US/10398221
 ; Publication No. US20040018514A1
 ; GENERAL INFORMATION:
 ; APPLICANT: KUNST, Frederik
 ; APPLICANT: GLASER, Philippe
 ; TITLE OF INVENTION: Listeria innocua, genome and applications
 ; FILE REFERENCE: 344 702 - US
 ; CURRENT APPLICATION NUMBER: US/10/398,221
 ; CURRENT FILING DATE: 2003-03-27
 ; PRIOR APPLICATION NUMBER: PCT/FR 01/03 061
 ; PRIOR FILING DATE: 2001-10-04
 ; PRIOR APPLICATION NUMBER: FR 00/12 697
 ; PRIOR FILING DATE: 2000-10-04
 ; NUMBER OF SEQ ID NOS: 4025
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 3163
 ; LENGTH: 1037
 ; TYPE: DNA
 ; ORGANISM: Listeria monocytogenes 4b
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (1)-(end)
 ; OTHER INFORMATION: n can be any nucleotide: a,g,c or t/u
 US-10-398-221-3163

Query Match 17.4%; Score 128; DB 16; Length 1037;
 Best Local Similarity 52.5%; Pred. No. 8.7e-17;
 Matches 355; Conservative 0; Mismatches 285; Indels 36; Gaps 2;
 Qy 25 ATACCCAGCCGAGCACTGTTTCAATCATTTTGTACGATCGTATGTTTAACTTGT 84
 Db 257 ATATCCAGACCCCAACTGTTTCAATCATTTCTAGCAGCGTCTTCTGATCATCATGCA 316
 Qy 85 TCGTAATATCTTATATAACATCTACGATTTGTTTCTACTTCACTA-----TTATA 138
 Db 317 TTTTTCGCGAAGCTTTTAAAGTGATCAATTTGTTCCGAATTCGCAATGCCAAGAATTCATA 376
 Qy 139 ACCCCATACATGATTTAAAAATTTGTTCCCGTTCGATAACATGTTTATTTTTCAGCTAG 198
 Db 377 ACGCCATCTCTTTTAAAGAGACGTCGCGTTCGATCTTTATCCGAGATCTCGCTAA 436
 Qy 199 AGATATAGTAATCATCTCTGTTTGTAAATTTCAATTTTTCGCGCATTTACCGTCAC 258
 Db 437 ATAATAAGCAAGTCATATCTTTTCGGTGTCAAAACCGATTTCTTTCCGTCGCAATAAC 496
 Qy 259 TTTAAAGCGTCTTATCAATTTGTAATCCGTTGACATCGATTAATATCTCTT----- 309
 Db 497 ACGTGAGCTTCGTTATCAATTTTCAATGTGGGAAGGTAATATGTCTCCCGGTGTTC 556
 Qy 310 -----TTGTGGCTGACGACGCTAAATTTGCAGGAATTTCTTGCTAA 348

Db 557 TCTGCTGATTTCTTCTGAGGATTTGTTTGGCGGACGACGAGCAAGTCTTTCACGCGAAGCAC 616
 Qy 349 AGTTTCTTCAATATCAACCGCTTAACTATATAATCTGTCGACCTAATCAAGCCAGC 408
 Db 617 AACTTCTCTTGGGCTAAATGGTTTACGATGTAATCATCTGCACCAACTTCAAAGCCTTG 676
 Qy 409 AACTTTGTCTATACGATATCACTTTTTCGCTGTAATTAATGATAGGTGTAGATTTGTTGTG 468
 Db 677 TACTCGGTAGCTTCTCTCTCTTTTTCAGTCAACATGACAAAGGTGTGATTTAAACTC 736
 Qy 469 TCTAATTTTTCGCAAAAATTTCTAAGCCATTAATGACGGCAACATTAATTAATATGAT 528
 Db 737 TCTCAGTTCAACGACAACTTCGATGCCATCTTTTACCAGGCATCATTAATCAAGTAGGAT 796
 Qy 529 TAAATCATAGTAATGGCTAAGCGCTTTATCTAAACCGTCTTGTCCATCATCTCTGTGC 588
 Db 797 TACTTCATAAATTAATGTTTTCAGCGCCATGCTTAATGCTTGCATCACCATCACTAGCTTCTTC 856
 Qy 589 CACATTGTAATTTTCAATGTGAGTTCCTCAATTCAGAAATCTTGTCTAAGTTTGTTCATC 648
 Db 857 AATACGATAGTTTCTCTTTTCAAGATACATCTTAAAGAGCGCTCGTATCGGTCTCTCA 916
 Qy 649 TTCTACTATTAATAAT 664
 Db 917 ATCCACAACAGCACT 932

RESULT 12
 US-10-398-221-886/c
 ; Sequence 886, Application US/10398221
 ; Publication No. US20040018514A1
 ; GENERAL INFORMATION:
 ; APPLICANT: KUNST, Frederik
 ; APPLICANT: GLASER, Philippe
 ; TITLE OF INVENTION: Listeria innocua, genome and applications
 ; FILE REFERENCE: 344 702 - US
 ; CURRENT APPLICATION NUMBER: US/10/398,221
 ; CURRENT FILING DATE: 2003-03-27
 ; PRIOR APPLICATION NUMBER: PCT/FR 01/03 061
 ; PRIOR FILING DATE: 2001-10-04
 ; PRIOR APPLICATION NUMBER: FR 00/12 697
 ; PRIOR FILING DATE: 2000-10-04
 ; NUMBER OF SEQ ID NOS: 4025
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 886
 ; LENGTH: 717
 ; TYPE: DNA
 ; ORGANISM: Listeria monocytogenes-EGD
 US-10-398-221-886

Query Match 17.2%; Score 126.4; DB 16; Length 717;
 Best Local Similarity 52.4%; Pred. No. 1.6e-16;
 Matches 354; Conservative 0; Mismatches 286; Indels 36; Gaps 2;
 Qy 25 ATACCCAGCCGAGCACTGTTTCAATCATTTTGTACGATCGTATGTTTAACTTGT 84
 Db 693 ATATCCAGACCCCAACTGTTTCAATCATTTCTCGCAGCGTCTTCTGATCATCATGCGAG 634
 Qy 85 TCGTAATATCTTATATAACATCTACGACATTTGTTTCTACTTCACTA-----TTATA 138
 Db 633 TTTTTCGCGCAGCTTTTAAAGTGATCAATTCGTTCCGAATTCGCAAGAAATTCATA 574
 Qy 139 ACCCCATACATGATTTAAAAATTTGTTCCCGTTCGATCAACATGTTTATTTTTCAGCTAG 198
 Db 573 ACGCCACACTCTTTTAAAGAGTACTCGCGGTTCGATCTTTTATCCGAGATTTAGCTAA 514
 Qy 199 AGATATAGTAATCATCTCTGTTTGTAAATTTCTTGTCAATTTTTCGCGCCATTTACCGTCAC 258
 Db 513 ATAATAAGCAAGTCATATCTTTTCGGTGTCAAAACCGATTTCTTTTCCATCGCAATAAC 454
 Qy 259 TTTAAAGCGTCTTATCAATTTGTAATACCGTTGACATCGATTAATATCTCTT----- 309
 Db 453 ACGTGAGCTTCGTTATCAATTTTAAATGTGGGAAGGTAATATGTCTCCCGGTGTGCC 394

```
QY 310 -----TTGTGGCTGACGACGTAAATTTGACGAATTTCTTGCTAA 348
Db 393 TCCTGCTGATTTCTCTGAGGATTTGTTGGCGGACGAAAGACAGCCTTTCACGCGAAGCAC 334
QY 349 AAGTTCTTCAATATCAAAAGGCTTAACATATATATGCTGTCGACCGTAAATCAAGCCGACG 408
Db 333 AACTTTCTCTTGGGCTAAATGGTTTACGATGTAATCATCTGCACCAACTTCAAGCCCTTG 274
QY 409 AACTTTGTCATACGCTATCACATTTTCGCTGTAATTAATATGATGATGATGATTTGTTG 468
Db 273 TACTCGGTAGTCTTCTCTTTTGGAGTCAACATGACGCGGTGGATTTAACTC 214
QY 469 TCTAATTTGGACAAATTTCTTAAGCCATTAATGACGCGCAACATTAATCTAATATGAT 528
Db 213 CCTCAGTTTCAGGACAACTTTCGATGCCATCTTTACGAGCATCATCAAAATCAAGTAGGAT 154
QY 529 TAAATCATAGTAAGCTTAAGCGCTTTATCAAAAGGCTTTGTCCTCATCATCTCTGTC 588
Db 153 TACTTCAATTAATTTGTTTCAGCGCCATGCTTAAATGCTTTCATCACCATCACTAGCTTCTTC 94
QY 589 CACATTTGTAATTTTCATGCTGAGTTTCAATTTCAAGAAATCTTGCTAAGTTTGTTCATC 648
Db 93 AATACGATGTTTCTTTCAGATACATCTTAAGAAAGGCTGCTATCCGGTCTCTCATC 34
QY 649 TTCTACTATTAATTT 664
Db 33 ATCCACAACAAGCACT 18
```

RESULT 13

US-09-070-927A-117
; Sequence 117, Application US/09070927A
; Patent No. US20020120116A1

GENERAL INFORMATION:

APPLICANT: Charles A. Kunsch
Patrick J. Dillon
Steven Barash
TITLE OF INVENTION: Enterococcus faecialis Polynucleotides and Polypeptides
NUMBER OF SEQUENCES: 982
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/070,927A
FILING DATE: 04-May-2000
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/046,655
FILING DATE: 1997-05-16
APPLICATION NUMBER: 60/044,031
FILING DATE: 1997-05-06
APPLICATION NUMBER: 60/066,009
FILING DATE: 1997-11-14

ATTORNEY/AGENT INFORMATION:

NAME: Kenley K. Hoover
REGISTRATION NUMBER: 40,302
REFERENCE/DOCKET NUMBER: PB369
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512

INFORMATION FOR SEQ ID NO: 117:

SEQUENCE CHARACTERISTICS:
LENGTH: 12022 base pairs

TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 117:
US-09-070-927A-117

Query Match 17.1%; Score 125.8; DB 9; Length 12022;
Best Local Similarity 53.8%; Pred. No. 6.4e-16;
Matches 365; Conservative 0; Mismatches 287; Indels 27; Gaps 4;

```
QY 20 ATCATATACCAACGCGACCAAGCTTTTCATCATTTTGTCCAGTCTGATGGTTT--- 76
Db 11128 ATCTATATCAACTCCCAACACTGTTTGAATGACTTTTCTCCACCAAGTGCCTTCA 11187
QY 77 AACTTGTGTTCCATAATATCTTATATAACATCTACGACATTTGTTCTACTTCACTATTA 136
Db 11188 ATCTATCTCTTAATGCTGACGCGAACCATCAGGGTTTGTAGCAGAAACAATCTTCT 11247
QY 137 TAACCCCATACATGATTTTAAATTTGTTCCGGTTGCATPAACATGGTTTATTTTACGCT 196
Db 11248 TGTGCCACACACGCTCAAAAATTTTCATCCGCACCTAAACACACGATTTGGATGACTAGCT 11307
QY 197 AGAAGATATAGTAATCATACTCTGTTTGTGTTTAAATTTCTGCGCCATTTACCG-- 254
Db 11308 AATAGATAGAGGATACCAAACTCAAGAGCGGTTAAATGAAATTTCTTCTCTTATTGTT 11367
QY 255 -TCACTTTAAAGCGTTCTTATCAATTTGTAATACCGTTGACATCGATAATATCTTTT-- 311
Db 11368 TTGACTTCGTGAGAGTCTTATTAATTAATTAATGCGCGGACTTCTAATTCATCGGGTCT 11427
QY 312 -----GTGCTGACGACGTAAATTTGCAAGAAATTTCTTGTCTAAAGTTCT 355
Db 11428 TCTTTGGTTAATTCATTTGGCTTCTACGTAAATAGACTTTTACCAGCCATTACTTCT 11487
QY 356 TCAATATCAAAACGCTTAACATATATATCGTCTGCACGGTAATCAAGCCGACCACTTTG 415
Db 11488 AAGGATTTAAAGGCTTTGGTGACATATCGTCTGCACCGCACTAAAGCCTTTGATTTTA 11547
QY 416 TCATACGTATCACATTTTTCGCTGTAATTAATATAGTAGTGTAGATTGTTGTCTTAAT 475
Db 11548 TCCATATCGGTGTTGTTTGTGTTTAAATTAATTAATTAATTTGCGGATCTTTTCGAAGT 11607
QY 476 TTGCGACAAAATTTCTAAGCCATTAATGACGCGCAACATTAATCTAATATGATTAATTC- 534
Db 11608 TCTTTCACTACTTCCATTCATCCATGATTTGGCATATAATGTCCAAGATTAATAAATCG 11667
QY 535 --ATAGTAATGGCTTAAGCGCTTATCTAAACCGCTTGTCCATCATACTCTGTGTCACA 592
Db 11668 ATATCTGGTGTGTGTGCAATTTGGATAACGCTTCTTTTCCATCAATAAGCTTTCCAGCT 11727
QY 593 TTGTAATTTTCATGTGTGAGTTCCAAATTTCAAGAAATCTTGTGAAGTTTGTTCATCTTCT 652
Db 11728 TCGTATCTTCAATATGAATAATAATACTCAATAGTTCTACAAATTTCTTTATCATCTCT 11787
QY 653 ACTATTAATAATTTGCGTCA 671
Db 11788 GCAACTAAATTTTTCATAA 11806
```

RESULT 14

US-09-925-637-45/c

; Sequence 45, Application US/09925637
; Patent No. US2002010338A1
GENERAL INFORMATION:
APPLICANT: Choi
TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Polypeptides
FILE REFERENCE: PB560
CURRENT APPLICATION NUMBER: US/09/925,637
CURRENT FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: PCT/US00/23773
PRIOR FILING DATE: 2000-08-31
PRIOR APPLICATION NUMBER: US 60/151,933
PRIOR FILING DATE: 1999-09-01

;; PRIOR APPLICATION NUMBER: US 08/781,986
;; PRIOR FILING DATE: 1997-01-03
;; PRIOR APPLICATION NUMBER: US 08/956,171
;; PRIOR FILING DATE: 1997-10-20
;; PRIOR APPLICATION NUMBER: US 60/009,861
;; PRIOR FILING DATE: 1996-01-06
;; NUMBER OF SEQ ID NOS: 74
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 45
;; LENGTH: 699
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-09-925-637-45

Query Match 17.1%; Score 125.6; DB 9; Length 699;
Best Local Similarity 52.3%; Pred. No. 2.4e-16;
Matches 324; Conservative 0; Mismatches 269; Indels 27; Gaps 1;
QY 73 TTTTAACTTCTGTTTCTGTAATAATCTTATATAAATCTTACGACATTTGTTTCTACTTCACT 132
DB 630 TTCAATCTTTTTCAGTAAACGACGAATCGTTACATCGCGTAGCTACATCGGCAAGTA 571
QY 133 ATTATAACCCCAATGATTTTAAATTTGTTCCCGTTGCATTAACATGTTTATTTTC 192
DB 570 ATCATAGCCCACTACTGTTTGTATAAATGTTTCAAGTGTCTTATTAATTTTCCCATATGTTT 511
QY 193 AGCTAGAGATATAGTAATATCATCTCTGTTTGTATTAATCAATTTCTGCGCATTTAC 252
DB 510 TGATAAATTAATGGAACATTTCAATTCAGATGTGTTAATTCANATATCTTCCGACGTTT 451
QY 253 CGTCACCTTTAAAGCGTTCTTATCAATTTGTAATCCGTTGACATCGATTAATATCCCTTTTG 312
DB 450 TTTAATAGATATGCGTCTGGATAAATCAATATCTTTAATTTGATTTTCATTCGTTAC 391
QY 313 TGGC-----TGACGACGTAATAATTCAGCAATTTCTTGC 345
DB 390 ATTTCCAGTGTCTGTGCTGGTTGTGAGTAATGACGACGTAAGTTGCTTTTCACAGTGC 331
QY 346 TAAAGTTCTTCAATATCAACCGCTTAATCTATATAATCGTCTGCACCGTAATCAAGCC 405
DB 330 GATTAAATTCACGGTACTTAACGGTTTCGTTACATAGTCACTCTGCACCTAGTTCTAAACC 271
QY 406 AGCAACTTTTGCATAGTATCATCTTTGCTGTAATATAATGATAGGTGTAGATTGTTG 465
DB 270 AAGCACTTTTATCAATTTCTGAATCTTTAGCAGTAAGCATTTATTTGGCATTTGTTATT 211
QY 466 TTGCTAATTTTGGCAAAATTTCTAAGCCATTAATTTGACGGCAACATTAATCTAATAT 525
DB 210 TTTGCGCACTTCACGACATCTTCCATACCACGACGAGGTAACATGATATCTAGTAA 151
QY 526 GATTAAATCATAGTAATGGCTAAGCGCTTTATCTAAACCGCTTGTGCTCACTACTCTGT 585
DB 150 TAGATGTCGTGTTCTTCTTCAATAATTAAGTGTGCTGATCATATTACCATCGTATGCACA 91
QY 586 GTCCACATTTGTAATTTTTCATGTGTAGTTCCAAATTCAGAAATCTTGTGTAAGTTTGTTC 645
DB 90 GTACACATCGTATCTCTTTTAAAGTTAAATTTCTAAATATATCAGCAATCGGTTTTC 31
QY 646 ATCTTCTACTATTAATAATTT 665
DB 30 ATCATCAACTACAACAATTT 11

RESULT 15
US-10-084-205-45/c
; Sequence 45, Application US/10084205
; Publication No. US20030049648A1
; GENERAL INFORMATION:
; APPLICANT: Choi, Gil
; TITLE OF INVENTION: 37 Staphylococcus aureus Genes and Polypeptides
; FILE REFERENCE: PB515P1
; CURRENT APPLICATION NUMBER: US/10/084,205
; CURRENT FILING DATE: 2002-02-28

;; PRIOR APPLICATION NUMBER: PCT/US00/23773
;; PRIOR FILING DATE: 2000-08-31
;; PRIOR APPLICATION NUMBER: 60/151,933
;; PRIOR FILING DATE: 1999-09-01
;; NUMBER OF SEQ ID NOS: 74
;; SOFTWARE: PatentIn Ver. 3.1
;; SEQ ID NO 45
;; LENGTH: 699
;; TYPE: DNA
;; ORGANISM: Staphylococcus aureus
US-10-084-205-45

Query Match 17.1%; Score 125.6; DB 15; Length 699;
Best Local Similarity 52.3%; Pred. No. 2.4e-16;
Matches 324; Conservative 0; Mismatches 269; Indels 27; Gaps 1;
QY 73 TTTTAACTTGTTCGTAAATATCTTATATAAATCTTACGACATTTGTTTCTACTTCACT 132
DB 630 TTCAATCTTTTTCAGTAAACGACGAATCGTTACATCGCGTAGCTACATCGGCAAGTA 571
QY 133 ATTATAACCCCAATGATTTTAAATTTGTTCCCGTTGCATTAACATGTTTATTTTC 192
DB 570 ATCATAGCCCACTACTGTTTGTATAAATGTTTCAAGTGTCTTATTAATTTTCCCATATGTTT 511
QY 193 AGCTAGAGATATAGTAATATCATCTCTGTTTGTATTAATCAATTTCTGCGCATTTAC 252
DB 510 TGATAAATTAATGGAACATTTCAATTCAGATGTGTTAATTCANATATCTTCCGACGTTT 451
QY 253 CGTCACCTTTAAAGCGTTCTTATCAATTTGTAATCCGTTGACATCGATTAATATCCCTTTTG 312
DB 450 TTTAATAGATATGCGTCTGGATAAATCAATATCTTTAATTTGATTTTCATTCGTTAC 391
QY 313 TGGC-----TGACGACGTAATAATTCAGCAATTTCTTGC 345
DB 390 ATTTCCAGTGTCTGTGCTGGTTGTGAGTAATGACGACGTAAGTTGCTTTTCACAGTGC 331
QY 346 TAAAGTTCTTCAATATCAACCGCTTAATCTATATAATCGTCTGCACCGTAATCAAGCC 405
DB 330 GATTAAATTCACGGTACTTAACGGTTTCGTTACATAGTCACTCTGCACCTAGTTCTAAACC 271
QY 406 AGCAACTTTTGCATAGTATCATCTTTGCTGTAATATAATGATAGGTGTAGATTGTTG 465
DB 270 AAGCACTTTTATCAATTTCTGAATCTTTAGCAGTAAGCATTTATTTGGCATTTGTTATT 211
QY 466 TTGCTAATTTTGGCAAAATTTCTAAGCCATTAATTTGACGGCAACATTAATCTAATAT 525
DB 210 TTTGCGCACTTCACGACATCTTCCATACCACGACGAGGTAACATGATATCTAGTAA 151
QY 526 GATTAAATCATAGTAATGGCTAAGCGCTTTATCTAAACCGCTTGTGCTCACTACTCTGT 585
DB 150 TAGATGTCGTGTTCTTCTTCAATAATTAAGTGTGCTGATCATATTACCATCGTATGCACA 91
QY 586 GTCCACATTTGTAATTTTTCATGTGTAGTTCCAAATTCAGAAATCTTGTGTAAGTTTGTTC 645
DB 90 GTACACATCGTATCTCTTTTAAAGTTAAATTTCTAAATATATCAGCAATCGGTTTTC 31
QY 646 ATCTTCTACTATTAATAATTT 665
DB 30 ATCATCAACTACAACAATTT 11

Search completed: October 5, 2004, 13:14:09
Job time : 383.125 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 5, 2004, 07:05:52 ; Search time 1100.87 Seconds
(without alignments)
10135.683 Million cell updates/sec

Title: US-09-006-627-1
Perfect score: 2201
Sequence: 1 TAAATTAAAGCAACTATT.....AAAACAAATAACAGTGCTT 2201

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 3340653 seqs, 2534783454 residues
Total number of hits satisfying chosen parameters: 6681306

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 100%
Listing first 45 summaries

Database : Published Applications NA.*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
18: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
19: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1432	65.1	1440	8	US-08-781-986A-725
2	1432	65.1	1440	13	US-10-329-624-725
3	1356	61.6	1356	13	US-10-282-122A-34148
4	673.2	30.6	4858	8	US-08-781-986A-379
5	673.2	30.6	4858	13	US-10-329-624-379
6	399.8	18.2	495269	16	US-10-398-221-8
7	399.8	18.2	3011208	16	US-10-398-221-2058
8	246.6	11.2	1452	13	US-10-282-122A-24752
9	197.4	9.0	1551	9	US-09-070-927A-472
10	194.8	8.9	1908	9	US-09-070-927A-27
11	154.8	7.0	684707	16	US-10-398-221-9
12	154.8	7.0	3011208	16	US-10-398-221-2058
13	141	6.4	1449	13	US-10-282-122A-15787
14	136.8	6.2	1839	13	US-10-282-122A-9474

15	133.4	6.1	1755	13	US-10-282-122A-35146	Sequence 35146, A
16	131.8	6.0	690	13	US-10-282-122A-17336	Sequence 17336, A
17	130.2	5.9	760	8	US-08-781-986A-822	Sequence 822, App
18	130.2	5.9	760	13	US-10-329-624-822	Sequence 822, App
c 19	130	5.9	31517	9	US-09-070-927A-180	Sequence 180, App
c 20	128	5.8	1037	16	US-10-398-221-3163	Sequence 3163, App
21	128	5.8	1707	13	US-10-282-122A-16047	Sequence 16047, A
22	127.6	5.8	9425	8	US-08-781-986A-87	Sequence 87, Appl
23	127.6	5.8	9425	13	US-10-329-624-87	Sequence 87, Appl
24	126.8	5.8	1737	9	US-09-815-242-4265	Sequence 4265, App
25	126.8	5.8	1752	9	US-09-815-242-8499	Sequence 8499, App
26	126.4	5.7	717	16	US-10-398-221-886	Sequence 886, App
c 27	125.8	5.7	12022	9	US-09-070-927A-117	Sequence 117, App
28	125.6	5.7	699	9	US-09-925-637-45	Sequence 45, Appl
29	125.6	5.7	699	15	US-10-084-205-45	Sequence 45, Appl
30	125.6	5.7	699	17	US-10-712-713-45	Sequence 45, Appl
31	125.6	5.7	708	9	US-09-815-242-8902	Sequence 8902, App
32	125.6	5.7	708	13	US-10-282-122A-8093	Sequence 8093, App
33	125	5.7	721	13	US-10-282-122A-34637	Sequence 34637, A
34	124.6	5.7	705	13	US-10-282-122A-10198	Sequence 10198, A
35	124.4	5.7	696	13	US-10-282-122A-16259	Sequence 16259, A
36	124	5.6	699	9	US-09-815-242-4621	Sequence 4621, App
37	124	5.6	708	9	US-09-815-242-8576	Sequence 8576, App
38	123.6	5.6	1767	13	US-10-282-122A-8000	Sequence 8000, App
39	123.6	5.6	2244	9	US-09-866-292-1	Sequence 1, Appl
40	122.8	5.6	319630	16	US-10-398-221-7	Sequence 7, Appl
41	122.2	5.6	678	9	US-09-815-242-7381	Sequence 7381, App
42	122.2	5.6	678	13	US-10-335-977-2707	Sequence 2707, App
43	122.2	5.6	678	13	US-10-335-977-2708	Sequence 2708, App
44	120	5.5	663	16	US-10-398-221-732	Sequence 732, App
45	120	5.5	663	16	US-10-398-221-2755	Sequence 2755, App

ALIGNMENTS

RESULT 1

US-08-781-986A-725
; Sequence 725, Application US/08781986A
; Publication No. US20030054436A1
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5255
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781,986A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Benson, Bob
; REGISTRATION NUMBER: 30,446
; REFERENCE/DOCKET NUMBER: PB248PP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 725:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1440 base pairs

```
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-781-986A-725

Query Match      65.1%; Score 1432; DB 8; Length 1440;
Best Local Similarity 99.7%; Pred. No. 2.4e-241;
Matches 1432; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 29 AAATACAAAATTTAAATATATTACCTTATTAGAAAAGTCGTAATATGAGGTGTACAAA 88
DB 4 AAATNCANTATTNAAATATATTACCTTATTAGAAAAGTCGTAATATGAGGTGTACAAA 63
QY 89 TGACGCAAAATTTAAATAGTAGAAGATGAACAAAATCTTAGCAAGATTTCTTGAATTCGAAC 148
DB 64 TGACGCAAAATTTAAATAGTAGAAGATGAACAAAATCTTAGCAAGATTTCTTGAATTCGAAC 123
QY 149 TCACACATGAAAATTTAAATGTGGACACAGAGTATGATGACAAGACGGTTTATAGATAAAG 208
DB 124 TCACACATGAAAATTTAAATGTGGACACAGAGTATGATGACAAGACGGTTTATAGATAAAG 183
QY 209 CGCTTAGCCATTACTATGATTTAATCATATTTAGATTTAATGTTGCCGTCAATTAATGGCT 268
DB 184 CGCTTAGCCATTACTATGATTTAATCATATTTAGATTTAATGTTGCCGTCAATTAATGGCT 243
QY 269 TAGAAAATTTCTGCAAAATTTAGACAACAATCTACACCTATCATTTATATATTTACAGCGA 328
DB 244 TAGAAAATTTCTGCAAAATTTAGACAACAATCTACACCTATCATTTATATTTACAGCGA 303
QY 329 AAAAGTGATACGTATGACAAAATTTGCTGGGTTGATTAACGGTCGAGACGATTTATATATTA 388
DB 304 AAAAGTGATACGTATGACAAAATTTGCTGGGTTGATTAACGGTCGAGACGATTTATATATTA 363
QY 389 AGCGTTTGTATATGAAGAACTTTTAGCAAGAAATTCGTGCAATTTTACGTCGTACGCCAC 448
DB 364 AGCGTTTGTATATGAAGAACTTTTAGCAAGAAATTCGTGCAATTTTACGTCGTACGCCAC 423
QY 449 AAAAGGATATTATTCGATGTCAACGGTATTCAATTTGATTAAGAACGCTTTTAAAGTGACGG 508
DB 424 AAAAGGATATTATTCGATGTCAACGGTATTCAATTTGATTAAGAACGCTTTTAAAGTGACGG 483
QY 509 TAAATGCGCAGAAAATTTGAATTTAAACAAAACAGAGTATGATTTACTATATCTTTCTAGCTG 568
DB 484 TAAATGCGCAGAAAATTTGAATTTAAACAAAACAGAGTATGATTTACTATATCTTTCTAGCTG 543
QY 569 AAAATAAAAACCATGTTATGCAACGGGAACAAAATTTTAAATCATGATGCGGGTTATATTA 628
DB 544 AAAATAAAAACCATGTTATGCAACGGGAACAAAATTTTAAATCATGATGCGGGTTATATTA 603
QY 629 GTGAAGTAGAAAACAAATGTCGTAGATGTTTATATAGATATTTTACGAAACAAAGTTAAAC 688
DB 604 GTGAAGTAGAAAACAAATGTCGTAGATGTTTATATAGATATTTTACGAAACAAAGTTAAAC 663
QY 689 CATACGATCGTGACAAAATGATTTGAAACAGTTCGTGGCGTTGGGTATGTGATACGATGAC 748
DB 664 CATACGATCGTGACAAAATGATTTGAAACAGTTCGTGGCGTTGGGTATGTGATACGATGAC 723
QY 749 AAAACGTAATTTGCGCAATTAATCTGGATTTATGTTTACCAAGATGATTTACGTTTGTACGAT 808
DB 724 AAAACGTAATTTGCGCAATTAATCTGGATTTATGTTTACCAAGATGATTTACGTTTGTACGAT 783
QY 809 ATTTTGTGTTTGTAAATTAATTTTCTTGTGAAGATACACTGCATAATAGTAGCT 868
DB 784 ATTTTGTGTTTGTAAATTAATTTTCTTGTGAAGATACACTGCATAATAGTAGCT 843
QY 869 TGATGATGCAAGACGAGCTCAAGCGATATTAATAATTTTCAATTTCAAGCGCTGTAA 928
DB 844 TGATGATGCAAGACGAGCTCAAGCGATATTAATAATTTTCAATTTCAAGCGCTGTAA 903
QY 929 AGATATATCTGCATTAGACTGAATGCATCTTTAGGTAATTTTCAAGAGATAATATTATTA 988
DB 904 AGATATATCTGCATTAGACTGAATGCATCTTTAGGTAATTTTCAAGAGATAATATTATTA 963

RESULT 2
US-10-329-624-725
; Sequence 725, Application US/10329624
; Publication No. US20040043037A1
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; Gil H. Choi
; Patrick S. Dillon
; Craig A. Rosen
; Steven C. Barash
; Michael R. Rannon
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5256
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: Hp Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/329,624
; FILING DATE: 27-Dec-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/956,171
; FILING DATE: October 20, 1997
; APPLICATION NUMBER: 60/009,861
; FILING DATE: January 5, 1996
; APPLICATION NUMBER: 08/781,986
; FILING DATE: January 3, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mark J. Hymen
; REGISTRATION NUMBER: 46,789
; REFERENCE/DOCKET NUMBER: PB248PID1
; TELECOMMUNICATION INFORMATION:
```

TELEPHONE: (240) 314-1224
TELEFAX: (301) 309-8439
INFORMATION FOR SEQ ID NO: 725:
SEQUENCE CHARACTERISTICS:
LENGTH: 1440 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO:
US-10-329-624-725

Query Match	Best Local Similarity	65.1%	Score 1432	DB 13	Length 1440
Matches 1432	Conservative 1	Mismatches 4	Indels 0	Gaps 0	
Qy	29	AAATACAAATTTTAAATATATATACCTTATTAGAAAAGTCGTATATAGCGGTGTACAA	88		
Db	4	AAATNCANTATTTNAAATATATATACCTTATTAGAAAAAGTCGTATATAGCGGTGTACAA	63		
Qy	89	TGACGCAATTTTAAATAGTAGAAGATGAACAAAACCTTAGCAAGATTCTTGTGAATTTGGAAC	148		
Db	64	TGACGCAATTTTAAATAGTAGAAGATGAACAAAACCTTAGCAAGATTCTTGTGAATTTGGAAC	123		
Qy	149	TCACACATGAAAATTTAACAATGTGGACACAGAGTAGATGAGCAAGAAGCGTTTGTAGATAAG	208		
Db	124	TCACACATGAAAATTTAACAATGTGGACACAGAGTAGATGAGCAAGAAGCGTTTGTAGATAAG	183		
Qy	209	CGCTTAGCCATTTACTATGATTTAATCATATTAGATTTAATGTTGGCGTCAATTAATGGCT	268		
Db	184	CGCTTAGCCATTTACTATGATTTAATCATATTAGATTTAATGTTGGCGTCAATTAATGGCT	243		
Qy	269	TAGAAATTTGTGCGAAATTTAGACAAACAACTCTACACCTTATCATTAATTAATTCACGGA	328		
Db	244	TAGAAATTTGTGCGAAATTTAGACAAACAACTCTACACCTTATCATTAATTAATTCACGGA	303		
Qy	329	AAAGTGATACGTATGACAAAAGTTGCTGGCGTTGATTACGGTCGACAGCATATATAGTTA	388		
Db	304	AAAGTGATACGTATGACAAAAGTTGCTGGCGTTGATTACGGTCGACAGCATATATAGTTA	363		
Qy	389	AGCGTTTGATATTGAAGAACTTTTGAAGAATTTCTGTCGAATTTTACGTCGTAGCCAC	448		
Db	364	AGCGTTTGATATTGAAGAACTTTTGAAGAATTTCTGTCGAATTTTACGTCGTAGCCAC	423		
Qy	449	AAAAGGATATTATCGATGTCACCGTATTACAAATTGATGAAGAACGCTTTTAAAGTGACGG	508		
Db	424	AAAAGGATATTATCGATGTCACCGTATTACAAATTGATGAAGAACGCTTTTAAAGTGACGG	483		
Qy	509	TAAATGGCGCAGAAATTCGAATTTACAAAACAGAGTAGATTTACTATATCTCTTAGCTG	568		
Db	484	TAAATGGCGCAGAAATTCGAATTTACAAAACAGAGTAGATTTACTATATCTCTTAGCTG	543		
Qy	569	AAAAATAAAAACCATGTTATGCAACGGGAACAAATTTTAAATCATGTATGGGTTTATAATA	628		
Db	544	AAAAATAAAAACCATGTTATGCAACGGGAACAAATTTTAAATCATGTATGGGTTTATAATA	603		
Qy	629	GTGAAGTAGAAACAAATGTCGTAGATGTTTATATAAGATATTTACGAAACAAAGTTTAAAC	688		
Db	604	GTGAAGTAGAAACAAATGTCGTAGATGTTTATATAAGRTATTTACGAAACAAAGTTTAAAC	663		
Qy	689	CATACGATCGTGACAAAATGATTGAAACAGTTCTGTGGCGTTGGGTATGTGATACATGAC	748		
Db	664	CATACGATCGTGACAAAATGATTGAAACAGTTCTGTGGCGTTGGGTATGTGATACATGAC	723		
Qy	749	AAAAACGTAAATTTGCGCAATAACTCGATTTATGTTTACACGATGATTAGCTTTGTCCAGAT	808		
Db	724	AAAAACGTAAATTTGCGCAATAACTCGATTTATGTTTACACGATGATTAGCTTTGTCCAGAT	783		
Qy	809	ATTTTTGTTTTGTTTTAATTTATTTTTTTCTTGAAAGATACACGTCAATATAGTGAGCT	868		
Db	784	ATTTTTGTTTTGTTTTAATTTATTTTTTTCTTGAAAGATACACGTCAATATAGTGAGCT	843		
Qy	869	TGATGATGCGAGAACGAAAGCTCAAGCGATTAATTAATTTATTTTCAATTTTAAAGCTGTTAA	928		

RESULT 3
US-10-282-122A-34148
; Sequence 34148, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625

TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 379:
SEQUENCE CHARACTERISTICS:
LENGTH: 4858 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear

US-08-781-986A-379

Query Match 30.6%; Score 673.2; DB 8; Length 4858;
Best Local Similarity 99.6%; Pred. No. 4.2e-108;
Matches 675; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1524 TGGGAAAAAGACCCAGCAGTATTAGAAGATCGTTAAATATTTCTATTGAAGAAATG 1583
DB 1 TGGGAAAAAGACCCAGCAGTATTAGAAGATCGTTAAATATTTCTATTGAAGAAATG 60

QY 1584 AATCGTATCATAAAATTAGTCGAAGATTAATGAAATGACTAAGAGAGATGATAATGC 1643
DB 61 AATCGTATCATAAAATTAGTCGAAGATTAATGAAATGACTAAGAGAGATGATAATGC 120

QY 1644 ATTTCTTCTGAAGCGCACGCCGTGCATATTATGATGAAATTCGCTCGCAATACACTCA 1703
DB 121 ATTTCTTCTGAAGCGCACGCCGTGCATATTATGATGAAATTCGCTCGCAATACACTCA 180

QY 1704 TTAAAAACAATTGCATCCTGATTATCAATTTTGATAGGATCTGACATCTAGAA 1763
DB 181 TTAAAAACAATTGCATCCTGATTATCAATTTTGATAGGATCTGACATCTAGAA 240

QY 1764 ATTTCTTCTGAAGCGCACGCCGTGCATATTATGATGAAATTCGCTCGCAATACACTCA 1823
DB 241 ATTTCTTCTGAAGCGCACGCCGTGCATATTATGATGAAATTCGCTCGCAATACACTCA 300

QY 1824 AAATATGATGCAAGAAATAAGAAATTAAGTTAAGTAAGAGTTAAAAATAAGCAAAA 1883
DB 301 AAATATGATGCAAGAAATAAGAAATTAAGTTAAGTAAGAGTTAAAAATAAGCAAAA 360

QY 1884 ATAATTGAATTTACAGATCATGGAATTCGTAATTTTAAATCTTTTATGATTAATGCAATC 1943
DB 361 ATAATTGAATTTACAGATCATGGAATTCGTAATTTTAAATCTTTTATGATTAATGCAATC 420

QY 1944 GATCGCTTTTATCGAGTGGATTAATCTCGTTCAAGAGTCAAGCGGTGATGACATCGGA 2003
DB 421 GATCGCTTTTATCGAGTGGATTAATCTCGTTCAAGAGTCAAGCGGTGATGACATCGGA 480

QY 2004 TTATCTATTGCTCAAAAAATCATTTCAATTTAAACGAGGATCGATTTAAAAATGAA 2063
DB 481 TTATCTATTGCTCAAAAAATCATTTCAATTTAAACGAGGATCGATTTAAAAATGAA 540

QY 2064 ATTAATAAGGAACAACGTTTAAAAATCATATTTTAAATCATGTCTGAGACGTCAATCAAAG 2123
DB 541 ATTAATAAGGAACAACGTTTAAAAATCATATTTTAAATCATGTCTGAGACGTCAATCAAAG 600

QY 2124 TCATAGATCAATTTTTTAAGTACACATTTAGCTGTGACTAATGATTAAGAACAACTATAA 2183
DB 601 TCATAGATCAATTTTTTAAGTACACATTTAGCTGTGACTAATGATTAAGAACAACTATAA 660

QY 2184 AACAAATAAACAGTGGTT 2201
DB 661 AACAAATAAACAGTGGTT 678

RESULT 5
US-10-329-624-379
Sequence 379, Application US/10329624
Publication No. US20040043037A1
GENERAL INFORMATION:
APPLICANT: Charles Kunsch
Gil H. Choi
Patrick S. Dillon
Craig A. Rosen
Steven C. Barash
Michael R. Fannon

TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
NUMBER OF SEQUENCES: 5256
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/329,624
FILING DATE: 27-Dec-2002
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/956,171
FILING DATE: October 20, 1997
APPLICATION NUMBER: 60/009,861
FILING DATE: January 5, 1996
APPLICATION NUMBER: 08/781,986
FILING DATE: January 3, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Mark J. Hyman
REGISTRATION NUMBER: 46,789
REFERENCE/DOCKET NUMBER: PB248PID1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (240) 314-1224
TELEFAX: (301) 309-8439
INFORMATION FOR SEQ ID NO: 379:
SEQUENCE CHARACTERISTICS:
LENGTH: 4858 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 379:
US-10-329-624-379

Query Match 30.6%; Score 673.2; DB 13; Length 4858;
Best Local Similarity 99.6%; Pred. No. 4.2e-108;
Matches 675; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1524 TGGGAAAAAGACCCAGCAGTATTAGAAGATCGTTAAATATTTCTATTGAAGAAATG 1583
DB 1 TGGGAAAAAGACCCAGCAGTATTAGAAGATCGTTAAATATTTCTATTGAAGAAATG 60

QY 1584 AATCGTATCATAAAATTAGTCGAAGATTAATGAAATGACTAAGAGAGATGATAATGC 1643
DB 61 AATCGTATCATAAAATTAGTCGAAGATTAATGAAATGACTAAGAGAGATGATAATGC 120

QY 1644 ATTTCTTCTGAAGCGCACGCCGTGCATATTATGATGAAATTCGCTCGCAATACACTCA 1703
DB 121 ATTTCTTCTGAAGCGCACGCCGTGCATATTATGATGAAATTCGCTCGCAATACACTCA 180

QY 1704 TTAAAAACAATTGCATCCTGATTATCAATTTTGATAGGATCTGACATCTAGAA 1763
DB 181 TTAAAAACAATTGCATCCTGATTATCAATTTTGATAGGATCTGACATCTAGAA 240

QY 1764 ATTTCTTCTGAAGCGCACGCCGTGCATATTATGATGAAATTCGCTCGCAATACACTCA 1823
DB 241 ATTTCTTCTGAAGCGCACGCCGTGCATATTATGATGAAATTCGCTCGCAATACACTCA 300

QY 1824 AAATATGATGCAAGAAATAAGAAATTAAGTTAAGTAAGAGTTAAAAATAAGCAAAA 1883
DB 301 AAATATGATGCAAGAAATAAGAAATTAAGTTAAGTAAGAGTTAAAAATAAGCAAAA 360

QY 1884 ATAATTGAATTTACAGATCATGGAATTCGTAATTTTAAATCTTTTATGATTAATGCAATC 1943
DB 361 ATAATTGAATTTACAGATCATGGAATTCGTAATTTTAAATCTTTTATGATTAATGCAATC 420

QY 1944 GATCGCTTTTATCGAGTGGATTAATCTCGTTCAAGAGTCAAGCGGTGATGACATCGGA 2003
DB 421 GATCGCTTTTATCGAGTGGATTAATCTCGTTCAAGAGTCAAGCGGTGATGACATCGGA 480

QY 2004 TTATCTATTGCTCAAAAAATCATTTCAATTTAAACGAGGATCGATTTAAAAATGAA 2063
DB 481 TTATCTATTGCTCAAAAAATCATTTCAATTTAAACGAGGATCGATTTAAAAATGAA 540

QY 2064 ATTAATAAGGAACAACGTTTAAAAATCATATTTTAAATCATGTCTGAGACGTCAATCAAAG 2123
DB 541 ATTAATAAGGAACAACGTTTAAAAATCATATTTTAAATCATGTCTGAGACGTCAATCAAAG 600

QY 2124 TCATAGATCAATTTTTTAAGTACACATTTAGCTGTGACTAATGATTAAGAACAACTATAA 2183
DB 601 TCATAGATCAATTTTTTAAGTACACATTTAGCTGTGACTAATGATTAAGAACAACTATAA 660

QY 2184 AACAAATAAACAGTGGTT 2201
DB 661 AACAAATAAACAGTGGTT 678

RESULT 5
US-10-329-624-379
Sequence 379, Application US/10329624
Publication No. US20040043037A1
GENERAL INFORMATION:
APPLICANT: Charles Kunsch
Gil H. Choi
Patrick S. Dillon
Craig A. Rosen
Steven C. Barash
Michael R. Fannon

RESULT 5
US-10-329-624-379
; Sequence 379, Application US/10329624
; Publication No. US20040040337A1
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; Gil H. Choi
; Patrick S. Dillon
; Craig A. Rosen
; Steven C. Barash
; Michael R. Fannon


```

Db 421 GATCGCTTTATCGAGTGGATAAATCTCGTTCAAGAGTCAAGCGGTAAATGGACTCGGA 480
Qy 2004 TTATCTATTGCTCAAAAATCATTAATAACGAGGATCGATTAATAATTAAGTGAA 2063
Db 481 TTATCTATTGCTCAAAAATCATTAATAACGAGGATCGATTAATAATTAAGTGAA 540
Qy 2064 ATTAATAAGGAACAACGCTTTAAATCATATTTTAATCATGCTCTGAGACGTCATCAAG 2123
Db 541 ATTAACAAGGAACAACGCTTTAAATCATATTTTAATCATGCTCTGAGACGTCATCAAG 600
Qy 2124 TCATAGGATCAATTTTAAAGTACACATTAATGCTGCTGATTAATGATTAAGAACACTATAA 2183
Db 601 TCATAGGATCAATTTTAAAGTACACATTAATGCTGCTGATTAATGATTAAGAACACTATAA 660
Qy 2184 AACAAATAAACAAGTGGTT 2201
Db 661 AACAAATAAACAAGTGGTT 678

RESULT 6
US-10-398-221-8/c
; Sequence 8, Application US/10398221
; Publication No. US20040018514A1
; GENERAL INFORMATION:
; APPLICANT: KUNST, Frederik
; APPLICANT: GLASER, Philippe
; TITLE OF INVENTION: Listeria innocua, genome and applications
; FILE REFERENCE: 344 702 - US
; CURRENT APPLICATION NUMBER: US/10/398, 221
; CURRENT FILING DATE: 2003-03-27
; PRIOR APPLICATION NUMBER: PCT/FR 01/03 061
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: FR 00/12 697
; PRIOR FILING DATE: 2000-10-04
; NUMBER OF SEQ ID NOS: 4025
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 8
; LENGTH: 495269
; TYPE: DNA
; ORGANISM: Listeria innocua
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(end)
; OTHER INFORMATION: n can be any nucleotide: a, g, c or t/u
US-10-398-221-8

Query Match 18.2%; Score 399.8; DB 16; Length 495269;
Best Local Similarity 52.8%; Pred. No. 2.1e-59;
Matches 1114; Conservative 0; Mismatches 902; Indels 93; Gaps 8;

Qy 85 CAAATGACGCAAAATTTAAATAGTAGAAGATGAACAAACCTTAGCAAGATTTCTTTGAATTG 144
Db 349529 CAAATGAATAGAAATTAATCGTAGAAGATGAACAAACCTTAGCAAGATTTCTTTGAAC 349470
Qy 145 GAACTCACACATGAATAATTAATGTGACACAGATGATGACAAAGCGTTTAGAT 204
Db 349469 GAATTAACAACGAAATTAATGAACCTGCTGTTGCTAATGATGACGCTGCTGGCTAGAA 349410
Qy 205 AAAGCGCTTAGCCATTTACTATGATTTAATCATATTAATGATTTAATGTTGCCGTCAATTAAT 264
Db 349409 CTCGCACTAAATGAAGATGGATGCTATTCTTACTAGACTTAATGTTTACCACATTTAAAC 349350
Qy 265 GGCTTAGAAATTTGTCGCAAAATTAGCAACACAACTACACCTATCATTTAATTAACA 324
Db 349349 GGGGTAGAGGTTTGCTGCTGCTGACGCAAGTGAACAAACACCCCATCATTTATGATAACT 349290
Qy 325 GCGAAAAGTGATAGCTATGACAAAGTTGCTGGGCTTCGATACGCTGACAGCGATTAATA 384
Db 349289 GCGGAGACTCTGTGATGATCGGTATCTGGACTTGACCGAGCAGATGATTAATC 349230
Qy 385 GTTAAGCCGTTGATATTGAAGAACTTTTAGCAAGAAATTCGTGCAATTTTACGTCGTCAG 444

```

```

Db 349229 GTTAAAGCCATTTGCAATTTGAAGAACTACTTGCAGCTCTCGGTTCACTATTACGCCGAGTA 349170
Qy 445 CCACAAAAGAT-----ATTATCGATGTCAACCGTATTACAAATTGAT 486
Db 349169 GAAATGTCAGAAACAATCTGCTAAACAAACACGCTACAATPACCGCAACTTAATCGTTGAA 349110
Qy 487 AAGAACGCTTTTAAAGTGCAGGTTAAATGGCGCAGAAATTCGAAATTAACAAAAACAGAGTAT 546
Db 349109 AAGAAAAATCGGATTTGAAACCGGATGAAGAAATTTATGATTTTAAACAAACCGGAGTAT 349050
Qy 547 GATTACTATATCTTCTAGCTGAAAAATAAAACCATGTTATGCAACGCGGAAACAAATTTTA 606
Db 349049 GAACTACTACTTCTTAAATGGAATAATGTCATATTTGTTCTTACTCGGGAAGTATTACTT 348990
Qy 607 AATCATGTATGGGTTTATATAGTAGAAGTAAGAAACAAATGTCGTAGATGTTTATATAAGA 666
Db 348989 ACAAAGTATGGGTTTATGAAACAGAAGTAGAAACAAATGTAGTAGATGATATATGTTCTG 348930
Qy 667 TATTTACGAAACCAAGTTTAAACCATACGATCGTGACAAATGATTGAAACAGTTCTGTGGC 726
Db 348929 TACTTGGAAATAAATTTGATCATCTCGAAGAAAGTTTACATCCAAACAGTTCCGCGG 348870
Qy 727 GTTGGGTATGTGATACG---ATGACAAAACGTAATTTGCGCAATAACTGATTAATTGTTA 783
Db 348869 ACTGGGTATGTGATCGTACATGACAACTAGCCCAATCTCTTAAAGTCTGTTCTTTGA 348810
Qy 784 CCAGGATGATTACGTTTGTGACGATATTTTGTGTTTAA-----TTATTAATTTTTT 838
Db 348809 AATTTAAATGGACTTTTGGAGCTAGTGCAGCGATTTTCTTAAACATTTTCTTATTTTCT 348750
Qy 839 CTTGAAAGATACACTGCATAATAGTCAGCTTGATGATGACGAGACGAGCTCAAGCCGATAT 898
Db 348749 ATCGGATTTATCAAGGATGCGCAATGTTGCTTAATGAAGAAACCGGAAGTAAAG 348690
Qy 899 TAAATAATTTATTTCTTAAGCCTGTTA-----AAGATATATCTGCAATTAGACTTGA 951
Db 348689 AACTGCTTCTAGCAACTACAGTACGTTAAACAAACCAAGATTTTAAACGCAATGAGAAA 348630
Qy 952 ATGCATCTTTAGTAAAT-----TTTCAAGAGATTAATTTATGATGAGCAATTAATAAA 1007
Db 348629 TCAATATTTTATTTAAATGAAGTAAACCGTTTAAATCGTAAATTAACAGATCAAGTGATTA 348570
Qy 1008 TTATTTGAGACATCGAATGATAACACAGTCAGAGTTG-----AACGAGTTA 1054
Db 348569 ATCTATACGATAAAGATGTCATTTTATTAAGTATTTATTTTCCAGAAACCAAGATA 348510
Qy 1055 TGAACACCGTTATTTTGACGCGTAAATAAAAAACGCTATAAAAGGCAATTGAATTAATTAAT 1114
Db 348509 TTACGAGTATCGATTTTTCGAGTATTTTGTAGCGCAGGATAAATTTATTTATGATA 348450
Qy 1115 TATTAAGAACCAATTTACACGCAAGAT-----1142
Db 348449 AGCCAAACCATTTGAAGGCGCAAAATGATGACTGCACAAAATGTCGCAAAATTTGGCAGATGATA 348390
Qy 1143 -----TTCAAAGGATAGCTTCTTAATTCATTCCTAGAAAATTTAGATAACATCG 1194
Db 348389 ATACGACAGTATCGGTTATGCGCAAGTGGTAAATCCGCTAACCTCTCTATAATCGGATGA 348330
Qy 1195 TAAATCATTTGATATATCATTTGCGCTGGCATTTGGAGTGAATTGCAACAAATTAATTAACCTGCCA 1254
Db 348329 TGGACCGGCTCTCTAGTAAACAAATTTGCTTGGAGCAGTAGCGCTCTTTATTTAGTGGAA 348270
Qy 1255 CAATCAGTTATGATTTTCAACACAAATTTACTTAACCGCTTGTGAGTTTATCAAAATAAA 1314
Db 348269 TGCCTCGTTACTTACTTAGCAGCAAGACTCTTAAATCCACTCACTCGGTTAGCAGCTACTA 348210
Qy 1315 TGATTCAGATTCGACGAGATGTTTTTCAAAATAAATTTGCAATTTAAATAACAAATTT---ATG 1371
Db 348209 TGAATGATTAATCGTAAATAATGTTTTTCAAAACGAATCGAAACGAACGAATTTCCAGAG 348150
Qy 1372 AAGAAATAGATAATTTAGCAAAATACGTTTAAATGAGATGATGAGCCAAATTTGAAGAAATCAT 1431
Db 348149 ATGAAATTCGCGAAATTCAGAGTTGTTTTTAAACGATATGATGACGCAATTTGAAACTAGTT 348090

```


Qy 1432 TTAATCAACAAAGACAATTTGTTGAAGATGCGTCAATGAATTAAGAACACCAATTAACAAA 1491
Db |||||
348089 TTGACGACAAAAACAATTCGTTGAAGATGCTTCCCATGAATCGGACACCGAGTTCAA 348030
Qy 1492 TTATTCAGGTCATTTAAATTTGATTACGAGATGGGGAAGAAAAGACCCAGCATATTAG 1551
Db |||||
348029 TCATGGAAAGGCACTTAAATTTACTAACTCGGTGGGGAAGAGACGCCCGCAGTACTGG 347970
Qy 1552 AAGAAATCGTTAAATTTCTATTGAAGAAATGAATCGTATCATATAAATTTAGTCGAAGAAT 1611
Db |||||
347969 ATGAATCGTTAAATGCTTCGTTAACGGAAATTAGAACGGATGAATAAATTTAGTACAAGAA 347910
Qy 1612 TACTTGAATTCAGTAAAGGAGATGAAATGACATTTCTTCAAGCGCAGACCGTGCATA 1671
Db |||||
347909 TGCTCGATTTATCAAGACGACAGACAGATTTCTCAACCGAAAGAAATTACAAATTTACCGATG 347850
Qy 1672 TTAATGATGAATTCGCTCGGATACACTCATTAACAATTTGATCTCTGATTTCAAT 1731
Db |||||
347849 TAAACGCAACAGTTGAGCAAGTGAGACGTAATTTTGAAGTAAATGATGAGAACTTTACGT 347790
Qy 1732 TTGATACGGATCTGACATCTAAATAATCTAGAAATTAATAATGAAATCTCATCAATTTCGAAC 1791
Db |||||
347789 TTACTTTAAAGAAAGATGATACGGATTTACGAGCGCTTTTCAACATATATTTAGAAC 347730
Qy 1792 AATTATTTTAACTTTTATGATTAATGCAATCAATATGATGTGAAGAAATTAAGAAATTA 1851
Db |||||
347729 AAATCTTAATTTATTTATGATTAACGCAATGAAATATTCGGGTGATGGCACTGAGGTGG 347670
Qy 1852 AAGTTAAGCAAGGTTAAATAATAGCAAAAATAATTAATTAATTAATTAATTAATTAATTA 1911
Db |||||
347669 ATATGCAATGCTATATAAGAAACAAAGCAAAATCCATATTTGATGTGCGGCAATTTACGAGAAG 347610
Qy 1912 GTATTCAGAGGAAGATCAAGATTTCAATTTTGTGATCGCTTTTATCGAGTGATTAATCTC 1971
Db |||||
347609 GCATCTCAAGAAAGAAATGATAGATCTTTAATCGTTTCTATCGTGTAGATAAGCCA 347550
Qy 1972 GTTCAAGAAATCAAGCGGTAATGAGATCGGATTTCTATTGCTCAAAAAATCAATTTCAAT 2031
Db |||||
347549 GAAGCGGTGAAAAAGGTGGTAATGCGCTCGCATTTGCTTAAACAATAGTTCGAAG 347490
Qy 2032 TAAACGGAGGATCGATTAATAATTAAGTGAATTAATAAGGAACAACGTTTAAATCA 2091
Db |||||
347489 GATATTTAGGAACGATTAATCGGTGAGTGAAGGATTAAGGATTAAGGATTAAGGATTA 347430
Qy 2092 TATTTTAAAT 2100
Db |||||
347429 CGCTTCAT 347421

RESULT 7

US-10-398-221-2058
; Sequence 2058, Application US/10398221
; Publication No. US20040018514A1
; GENERAL INFORMATION:
; APPLICANT: KUNST, Frederik
; TITLE OF INVENTION: Listeria innocua, genome and applications
; APPLICANT: GLASER, Philippe
; FILE REFERENCE: 344 702 - US
; CURRENT APPLICATION NUMBER: US/10/398,221
; CURRENT FILING DATE: 2003-03-27
; PRIOR APPLICATION NUMBER: PCT/FR 01/03 061
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: FR 00/12 697
; PRIOR FILING DATE: 2000-10-04
; NUMBER OF SEQ ID NOS: 4025
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 2058
; LENGTH: 3011208
; TYPE: DNA
; ORGANISM: Listeria innocua

US-10-398-221-2058

Query Match 18.2%; Score 399.8; DB 16; Length 3011208;
Best Local Similarity 52.8%; Pred. No. 4.3e-59;
Matches 1114; Conservative 0; Mismatches 902; Indels 93; Gaps 8;
Qy 85 CAAATGACGCAAAATTTAAATAGTAGAAGATGAAACAAACACTTAGCAAGATTTCTTGAATTTG 144
Db 1409422 CAAATGAATAGAAATTTAAATCGTAGAAGATGAAAAAATCTTAGCAGCGCTTTATTGAACATA 1409481
Qy 145 GAACTCACACATGAAATTTACAATGTGCACACAGATGATGACCAAGACGGTTTATAGT 204
Db 1409482 GAATTTACAACACGAAATTTATGAACCTGCTGCTTAATGATGACGCTGCGGTAGAA 1409541
Qy 205 AAAGCGTTAGCCATTTACTATGATTTAATCATATTAAGATTTAATTTGTTGCCGTCATTAAT 264
Db 1409542 CTCGCACTAAATGAAGATGGATGCTATTTTACTAGACTTAATGTTTACCAATTTAAAC 1409601
Qy 265 GGCCTTAGAAATTTCTGCCAAATTTAGACACAAACAACTACACCTATCATTTATATTAACA 324
Db 1409602 GGGTAGAGGTTTGTGCTGCTGATGACCAAGTGAACCAACACCCATCATTTATGATAACT 1409661
Qy 325 GCGAAAAATGATACGTATGACAAAGTTGCTGGGCTTGATTTACGGTGCAGACGATTAATA 384
Db 1409662 GCGGAGACTCTGATGATGATCGGTATCTGGACTTGACACGAGCAGATGATTAATC 1409721
Qy 385 GTTAAGCGTTGATATGTAAGAACTTTTACAGAAATTCGTGCAATTTTACGTCGTCAG 444
Db 1409722 GTTAAGCAATTTGCAATTTGAAGAACTACTTGCAGCTCTGCGTTCCTACTATTACGCGAGTA 1409781
Qy 445 CCACAAAAGNT-----ATTATCGATGTCAACGGTATTACAAATTTGAT 486
Db 1409782 GAAATCGAGAACAACTCTGCTAAACAAACACGCTACAATACCCCACTTAATCGTTGAA 1409841
Qy 487 AAGAACGCTTTTAAAGTGACGCTAAATGGCGCAGAAATTTGAATTTAAACAAAACAGAGTAT 546
Db 1409842 AAAGAAATCGGATTTGAAACGCGATGAAGAAATTTATGATTTTAAACAAACGCGAGTAT 1409901
Qy 547 GATTTACTATATCTTCTAGCTGAAATTAATAACCAATGTTATGCAACGCGAAACAAATTTTA 606
Db 1409902 GAACTACTACTTACTTTAAATGGAATAATGTCATATTTCTTCTACTCGGAAAGTATTACTT 1409961
Qy 607 AATCATGTATGGGTTTAAATAGTAGAAGTAAACAAATGCTCGTAGATGTTTATATAAGA 666
Db 1409962 AACAAAGTATGGGTTTATGAAACAGAGTGAACAAATGATGATGATGATGATGATGATGAT 1410021
Qy 667 TATTTAGCAACAAAGTTTAAACCAATACATCGTGCACAAAATGATTTGAAACAGTTTCGTGGC 726
Db 1410022 TACTTGGAAATATAAATGATCATCTCGACGAAAGTTACATCCAAACAGTTTCGCGGG 1410081
Qy 727 GTTGGGTATGTGATACG---ATGACAAAACGTTAAATTTGCGCAATTAACCTGGATTTATGTTA 783
Db 1410082 ACTGGGTATGTGATGCGTACATGACAACTAGCCCATTTCTCTTAAAGTCTGTTCTTTGA 1410141
Qy 784 CCAGATGATTACGTTTGTGTCAGATATTTTGTGTTGTTTAA-----TTATTTATTTT 838
Db 1410142 AATTTAAATGGACTTTTGGAGCTAGTGCAGCGATTTTCTTAAACATTTTCTTATTTTCTCT 1410201
Qy 839 CTTGAAAGATACACTGTCATAATAGTAGCTTTGATGTCAGAGACGAAAGCTCAACGCGATAT 898
Db 1410202 ATGCGATTTATCAAGGATTTGGGCAATGTTGCTTAATGAAGAAACCGGAAAGTAAAG 1410261
Qy 899 TAATAATTTATTTCTAAGCCTGTTA-----AAGATATATCTGCATTAGACTTTGA 951
Db 1410262 AACTGCTTCTAGCAACTACAAGTACGTTTAAACAAACCAAGATTTTAAACGCAATCAGGAAA 1410321
Qy 952 ATGCATCTTTAGTAAAT-----TTTCAAGAGATATATTTATGATGAGCATTAATAATAA 1007
Db 1410322 TCAATATTTTAAATACGATAAAACCGTTTATCGTAAATTTTAAAGATCAAGATGATTA 1410381
Qy 1008 TTATTTGAGACATCGAATGATAACACAGTGAAGTTG-----AACCAGGTTA 1054
Db 1410382 ATCTATACGATAAAGATGGTCAATTTTATTAATAGTATTTATTTTCCAGAAACCAAGATA 1410441
Qy 1055 TGAACACCGTTATTTTGACCGCGTAAATAAAAAACGCTATAAAGGCATTTGAATTTTAAAT 1114

Db 1410442 TTACGAGTATCGATTTTTCGACGATTTTGTAGCGCAGGATATAATTTATTATGAATA 1410501
QY 1115 TATTAAAGAACCAATTACACGCAAGAT-----1142
Db 1410502 AGCCAACCAATTGAAGGCGAATAATGATGACTGCACAAATGCCAATTTGTGGCAGATGATA 1410561
QY 1143 -----TTCAAAGGATATAGCTTGTAAATTCATTCACATAGAAATTTATGATAACATCG 1194
Db 1410562 ATAGGACAGTATCGGTTATGCGCAAGTGTAAATCGCTTAATCTCTTATTAATCGGATGA 1410621
QY 1195 TAAATCATTTGATATCATTTGCGTGGCAATTTGAGAGTGATGCAACAATTTATACTGCCA 1254
Db 1410622 TGGACCGGCTTCTAGTAAACAATGATTTTGTCTGGAGCAGTAGCGCTCTTTTATTAGTGGAA 1410681
QY 1255 CAATCAGTTATGATTTTCAACACAAATTAATCAACCGCTTGTCAGTTTATCAAAATAAA 1314
Db 1410682 TGCTCGGTTACTTACTAGCAGACAACTTCTTAAATCCACTCACTCGGTTAGCAGCTACTA 1410741
QY 1315 TGATTCAGATTCGACGAGATGGTTTTCAAAATAAATTTGCAATTAATAATACAAATTT---ATG 1371
Db 1410742 TGAATGATATCGTAAATATGGTTTTTCAAAACGAATCGAAACGAAACGAAATTCAGAG 1410801
QY 1372 AAGAAATPAGATAATTTAGCAAAATACGTTTAAATGAGATGATGAGCCAAATTTGAAGAATCAT 1431
Db 1410802 ATGAAATTTGGCAATTCACAGTTGTTTTTAAACGATATGATGACGCGAATTTGAAACTAGTT 1410861
QY 1432 TTAATCAACAAGACAAATTTTGAAGATCGGTCACATGAATTTACGACACCAATTCACAA 1491
Db 1410862 TTGAGCAGCAAAAACAAATTCGTTGAAGATCTTCCCATGAATTCGCGACACCAAGTTCAAA 1410921
QY 1492 TTATTCAGGTCATTTAAATTTGATTCAGCGATGGGGAAGAAAGACCCAGCAGTATTAG 1551
Db 1410922 TCAATGAGGCACTTAAATTTACTAATCTCGTGGGGAAGAGACCCGCGAGTACTGG 1410981
QY 1552 AAGAAATGTTAAATTTCTTATTGAAGAAATGAATCGTATCATATAAATTTAGTCAAGAAAT 1611
Db 1410982 ATGAATCGTTAAATGCTTCGTTAACGGAATTTAGAACGGATGAAATAATTTAGTCAAGAAA 1411041
QY 1612 TACTTGAATTTGACTAAAGGAGATGTAAATGACATTTCTTCTGAAGCGCAGACCGTGATGA 1671
Db 1411042 TGCTCGATTTTATCAAGACGAGAACAGATTTCTCAAAAGAAAGAAATTTACAAATTTACCGGATG 1411101
QY 1672 TTAATGATGAAATTCGCTCGCAATACACTCATTTAAACAAATTTGCATCTCTGATTTATCAAT 1731
Db 1411102 TAAACGCAAGTTGACCAAGTGAAGCTAATTTTGAAGTAATGTATGAGAACTTTTACGT 1411161
QY 1732 TTGATACGGATCTGCATCTTAAATAATCTAGAAATTTAAATGAAACCTTCATCAATTTGAAAC 1791
Db 1411162 TTACTTTAAAGAAGATGATACGGAATTTACGAGCGCTTATTCACATAATCATTTAGAAC 1411221
QY 1792 AATATTTTAAATCTTTTATGATATGCAATCAATATGATGATGAAGAAATTAAGAAATTA 1851
Db 1411222 AATCTTTAAATTTATTTATGATGATAACGCACTGAATATTCGGGTGATGGCAGCTGAGTGG 1411281
QY 1852 AAGTTAAGACAGGTTTAAATAAATAAGCAAAATAATTTGAAATTTACAGATCATGGAATTCG 1911
Db 1411282 ATATGATGCTATAAGAACAAAGCAATTCATATTTGATGTCGCGGATTCGAGAGAG 1411341
QY 1912 GTATTCAGAGGAGATCAAGATTTTCAATTTTGTATCGCTTTTATCGAGTGATGATAATCTC 1971
Db 1411342 GCATCTCACAGAAGAAATTTGATAAGATCTTTAATCGTTTCTATCGTGTAGATAAGGCCA 1411401
QY 1972 GTTCAAGAGTCAAGCGGTAATGGAATCGGATTCATTTATGCTCAAAAATCATTTCAAT 2031
Db 1411402 GAAGCGGTGAAAAGGTTGTAATGGCCCTCGGACTTGGGATTTGTAACAAATTTAGTCTGAAG 1411461
QY 2032 TAAACGGAGGATCGATTTAAATTTAAAGTCAAAATTAATAAGGAACAAACGCTTTAAATCA 2091
Db 1411462 GATATTTAGAACGATTTATGCGGTGAGTGGCGGATAAGGTACTACTATATAAATAATTA 1411521
QY 2092 TATTTTAAAT 2100
|||

Db 1411522 CGCTCCAT 1411530
RESULT 8
US-10-282-122A-24752
; Sequence 24752, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zvekund, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 24752
; LENGTH: 1452
; TYPE: DNA
; ORGANISM: Listeria monocytogenes
US-10-282-122A-24752

Query Match 11.2%; Score 246.6; DB 13; Length 1452;
Best Local Similarity 55.9%; Pred. No. 1.6e-33;
Matches 490; Conservative 0; Mismatches 384; Indels 3; Gaps 1;
QY 1223 ATTTGAGTGATTTGCAACAATTTATACTGCAACAATCAGTTATGTTATTTTCAACACAAT 1282
Db 549 ACTTGGGCGAGTGGCGCTCTTTATTAGTGAATGCTCGGCTACTTGTAGCACAATAATTT 608
QY 1283 TACTAAACCGCTTGTTCAGTTTATCAAAATAAATGATTGAGATTCGACGAGATCGTTTCA 1342
Db 609 CTTAAACCACTGACTCGTCTAGCTCGAACATGAATGATTTTCTGTAATAATCGTTTCA 668
QY 1343 AATAAATTCGATTTAAATACAAATTTAT---GAAGAAATAGATAAATTTAGCAAAATACGTT 1399
Db 669 AAAACGAATCGAGACGAAACAAATTTCTCGCGATGAAATTTGGCGAATTTGACCGTTGTTT 728
QY 1400 TAATGATGATGAGCGCAAAATTTGAAGAAATCAATTTTAAATCAACAAGAACAAATTTGTTGAGA 1459
Db 729 TAATGATGATGAGCGCAAAATTTGAAACTAGTTTTCGAACAGCAAAAGCAATTTTGTGAGGA 788
|||

```
QY 1460 TCGTCACATGAATTCAGAACACCAATACAAATTAATCAAGTCAATTAATAATTTGATTCA 1519
D 1461 TCGTCACATGAATTCAGAACACCAATACAAATTAATCAAGTCAATTAATAATTTGATTCA 1520
D 1462 TCGTCACATGAATTCAGAACACCAATACAAATTAATCAAGTCAATTAATAATTTGATTCA 1521
QY 1520 CGGATGGGGAAGAAAGACCCAGCAGTATTAGAACAAATCGTTAAATATTTCTATTGAAGA 1579
D 1521 CGGATGGGGAAGAAAGACCCAGCAGTATTAGAACAAATCGTTAAATATTTCTATTGAAGA 1580
D 1522 CGGATGGGGAAGAAAGACCCAGCAGTATTAGAACAAATCGTTAAATATTTCTATTGAAGA 1581
QY 1580 AATGAATCGTATCATATAAATAGTCCGAAGAATTAATGAAATTAATGAAAGAGATGTAAA 1639
D 1581 AATGAATCGTATCATATAAATAGTCCGAAGAATTAATGAAATTAATGAAAGAGATGTAAA 1640
D 1582 AATGAATCGTATCATATAAATAGTCCGAAGAATTAATGAAATTAATGAAAGAGATGTAAA 1641
QY 1640 TGACATTTCTTGAAGCGCAGACCGTGCATTAATGATGAATTAATGATGAATTAATGATGA 1699
D 1641 TGACATTTCTTGAAGCGCAGACCGTGCATTAATGATGAATTAATGATGAATTAATGATGA 1700
D 1642 TGACATTTCTTGAAGCGCAGACCGTGCATTAATGATGAATTAATGATGAATTAATGATGA 1701
QY 1700 CTCATTAAGCAATTCGATCTGATTAATCAATTTGATACGGATCTGACATCTTAAATCT 1759
D 1701 CTCATTAAGCAATTCGATCTGATTAATCAATTTGATACGGATCTGACATCTTAAATCT 1760
D 1702 CTCATTAAGCAATTCGATCTGATTAATCAATTTGATACGGATCTGACATCTTAAATCT 1761
QY 1760 AGAAATTAAGCAATTCGATCTGATTAATCAATTTGATACGGATCTGACATCTTAAATCT 1819
D 1761 AGAAATTAAGCAATTCGATCTGATTAATCAATTTGATACGGATCTGACATCTTAAATCT 1820
D 1762 AGAAATTAAGCAATTCGATCTGATTAATCAATTTGATACGGATCTGACATCTTAAATCT 1821
QY 1820 AATCAATTAAGCAATTCGATCTGATTAATCAATTTGATACGGATCTGACATCTTAAATCT 1879
D 1821 AATCAATTAAGCAATTCGATCTGATTAATCAATTTGATACGGATCTGACATCTTAAATCT 1880
D 1822 AATCAATTAAGCAATTCGATCTGATTAATCAATTTGATACGGATCTGACATCTTAAATCT 1881
QY 1880 AATCAATTAAGCAATTCGATCTGATTAATCAATTTGATACGGATCTGACATCTTAAATCT 1939
D 1881 AATCAATTAAGCAATTCGATCTGATTAATCAATTTGATACGGATCTGACATCTTAAATCT 1940
D 1882 AATCAATTAAGCAATTCGATCTGATTAATCAATTTGATACGGATCTGACATCTTAAATCT 1941
QY 1940 TTTTCATCGCTTTTTCGATGATTAATCAATTTGATACGGATCTGACATCTTAAATCT 1999
D 1941 TTTTCATCGCTTTTTCGATGATTAATCAATTTGATACGGATCTGACATCTTAAATCT 2000
D 1942 TTTTCATCGCTTTTTCGATGATTAATCAATTTGATACGGATCTGACATCTTAAATCT 2001
QY 2000 CGGATTAATTAAGCAATTCGATCTGATTAATCAATTTGATACGGATCTGACATCTTAAATCT 2059
D 2001 CGGATTAATTAAGCAATTCGATCTGATTAATCAATTTGATACGGATCTGACATCTTAAATCT 2060
D 2002 CGGATTAATTAAGCAATTCGATCTGATTAATCAATTTGATACGGATCTGACATCTTAAATCT 2061
QY 2060 TGAATTAATTAAGCAATTCGATCTGATTAATCAATTTGATACGGATCTGACATCTTAAATCT 2096
D 2061 TGAATTAATTAAGCAATTCGATCTGATTAATCAATTTGATACGGATCTGACATCTTAAATCT 2097
D 2062 TGAATTAATTAAGCAATTCGATCTGATTAATCAATTTGATACGGATCTGACATCTTAAATCT 2098
QY 2096 TGAATTAATTAAGCAATTCGATCTGATTAATCAATTTGATACGGATCTGACATCTTAAATCT 2145
D 2097 TGAATTAATTAAGCAATTCGATCTGATTAATCAATTTGATACGGATCTGACATCTTAAATCT 2146
D 2098 TGAATTAATTAAGCAATTCGATCTGATTAATCAATTTGATACGGATCTGACATCTTAAATCT 2147
```

RESULT 9

```
US-09-070-927A-472/c
; Sequence 472, Application US/09070927A
; Patent No. US20020120116A1
; GENERAL INFORMATION:
; APPLICANT: Charles A. Kunsch
;              Patrick J. Dillon
;              Steven Barash
; TITLE OF INVENTION: Enterococcus faecialis Polynucleotides and Polypeptides
; NUMBER OF SEQUENCES: 962
; CORRESPONDENCE ADDRESS:
; ADDRESS: Human Genome Sciences, Inc.
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/070,927A
; FILING DATE: 04-May-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/046,655
```

```
; FILING DATE: 1997-05-16
; APPLICATION NUMBER: 60/044,031
; FILING DATE: 1997-05-06
; APPLICATION NUMBER: 60/066,009
; FILING DATE: 1997-11-14
; ATTORNEY/AGENT INFORMATION:
; NAME: Kenley K. Hoover
; REGISTRATION NUMBER: 40,302
; REFERENCE/DOCKET NUMBER: PB369
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 472:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1551 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 472:
US-09-070-927A-472
```

```
Query Match          9.0%; Score 197.4; DB 9; Length 1551;
Best Local Similarity 51.6%; Pred. No. 7e-25;
Matches 494; Conservative 3; Mismatches 454; Indels 6; Gaps 2;

QY 1126 CAATTACAACGCAAGATTTCAAAGCGGTATAGCTTTGTTAAATTCATTCACTAGAAAATATG 1185
D 1127 CAATTACAACGCAAGATTTCAAAGCGGTATAGCTTTGTTAAATTCATTCACTAGAAAATATG 1186
QY 1186 ATACATCGTAAATCAATGTTATATCAATGCGCTGGCAATTTGGAGTGAATTCGAACAATTA 1245
D 1187 ATACATCGTAAATCAATGTTATATCAATGCGCTGGCAATTTGGAGTGAATTCGAACAATTA 1246
D 1188 ATACATCGTAAATCAATGTTATATCAATGCGCTGGCAATTTGGAGTGAATTCGAACAATTA 1247
QY 897 ATGAAATTCGTAATCAATTTATTAATTAACACATGTTAGTTTAAAGAGTATTTCTCTGATTG 838
D 898 ATGAAATTCGTAATCAATTTATTAATTAACACATGTTAGTTTAAAGAGTATTTCTCTGATTG 839
D 899 ATGAAATTCGTAATCAATTTATTAATTAACACATGTTAGTTTAAAGAGTATTTCTCTGATTG 840
QY 1246 TAATCGCCACAAATCAGTTATGTTATTTCAACACAAATTAATTAACCGCTTCTCAGTTTAT 1305
D 1247 TAATCGCCACAAATCAGTTATGTTATTTCAACACAAATTAATTAACCGCTTCTCAGTTTAT 1306
D 1248 TAATCGCCACAAATCAGTTATGTTATTTCAACACAAATTAATTAACCGCTTCTCAGTTTAT 1307
QY 837 TAATGTTAGTCTTATAGGCTTTATCTCTCTTCTTAAATTCATTAAAGAGTATTTGA 778
D 838 TAATGTTAGTCTTATAGGCTTTATCTCTCTTCTTAAATTCATTAAAGAGTATTTGA 779
D 839 TAATGTTAGTCTTATAGGCTTTATCTCTCTTCTTAAATTCATTAAAGAGTATTTGA 780
QY 1306 CAATTAATAATGATTGAGATTTCGACGAGATGTTTTCATAAAT---AAATTCGAAATTAATA 1362
D 1307 CAATTAATAATGATTGAGATTTCGACGAGATGTTTTCATAAAT---AAATTCGAAATTAATA 1363
D 1308 CAATTAATAATGATTGAGATTTCGACGAGATGTTTTCATAAAT---AAATTCGAAATTAATA 1364
QY 777 GAGACAGATGATACAAATCCGCAAGACCCGCAATCAGATGTCCACATCGCGGAGATTA 718
D 778 GAGACAGATGATACAAATCCGCAAGACCCGCAATCAGATGTCCACATCGCGGAGATTA 719
D 779 GAGACAGATGATACAAATCCGCAAGACCCGCAATCAGATGTCCACATCGCGGAGATTA 720
QY 1363 CAATTAATAATGATTGAGATTTCGACGAGATGTTTTCATAAAT---AAATTCGAAATTAATA 1422
D 1364 CAATTAATAATGATTGAGATTTCGACGAGATGTTTTCATAAAT---AAATTCGAAATTAATA 1423
D 1365 CAATTAATAATGATTGAGATTTCGACGAGATGTTTTCATAAAT---AAATTCGAAATTAATA 1424
QY 717 ATACGAGAGATGAGTTAGCAGATATCTCGGAAATCTTTAAATGAAATGTTAGATCGTATGA 658
D 718 ATACGAGAGATGAGTTAGCAGATATCTCGGAAATCTTTAAATGAAATGTTAGATCGTATGA 659
D 719 ATACGAGAGATGAGTTAGCAGATATCTCGGAAATCTTTAAATGAAATGTTAGATCGTATGA 660
QY 1423 AAGAAATCAATTAATAATGAGAAATTCGTTGAGATGCGTCAATGAAATTAATAAATTAATA 1482
D 1424 AAGAAATCAATTAATAATGAGAAATTCGTTGAGATGCGTCAATGAAATTAATAAATTAATA 1483
D 1425 AAGAAATCAATTAATAATGAGAAATTCGTTGAGATGCGTCAATGAAATTAATAAATTAATA 1484
QY 657 GACGCTATATCGAACAACAGAGCAGTTTGTGAAGATGTTTCCCATGAATTAAGAAGCAG 598
D 658 GACGCTATATCGAACAACAGAGCAGTTTGTGAAGATGTTTCCCATGAATTAAGAAGCAG 599
D 659 GACGCTATATCGAACAACAGAGCAGTTTGTGAAGATGTTTCCCATGAATTAAGAAGCAG 600
QY 1483 CAATTAATAATGATTGAGATTTCGACGAGATGTTTTCATAAAT---AAATTCGAAATTAATA 1542
D 1484 CAATTAATAATGATTGAGATTTCGACGAGATGTTTTCATAAAT---AAATTCGAAATTAATA 1543
D 1485 CAATTAATAATGATTGAGATTTCGACGAGATGTTTTCATAAAT---AAATTCGAAATTAATA 1544
QY 597 CCGTTGCGATTAATGAGAGGCAATTTAAACCTTTTAAATCGTTGGGGGAAAGACGATCTG 538
D 598 CCGTTGCGATTAATGAGAGGCAATTTAAACCTTTTAAATCGTTGGGGGAAAGACGATCTG 539
D 599 CCGTTGCGATTAATGAGAGGCAATTTAAACCTTTTAAATCGTTGGGGGAAAGACGATCTG 540
QY 1543 CAGTATTAGAAGAAATCGTTAAATATTTCTATTGAAGAAATGAATCGTATCATATAAATTAAG 1602
D 1544 CAGTATTAGAAGAAATCGTTAAATATTTCTATTGAAGAAATGAATCGTATCATATAAATTAAG 1603
D 1545 CAGTATTAGAAGAAATCGTTAAATATTTCTATTGAAGAAATGAATCGTATCATATAAATTAAG 1604
QY 537 AAATTTTAGATGAATCATTTAAAGGCCAGTTTACAAAGAAATTTAGTCGTATGAAGAGTTGG 478
D 538 AAATTTTAGATGAATCATTTAAAGGCCAGTTTACAAAGAAATTTAGTCGTATGAAGAGTTGG 479
D 539 AAATTTTAGATGAATCATTTAAAGGCCAGTTTACAAAGAAATTTAGTCGTATGAAGAGTTGG 480
QY 1603 TCGAAGAAATTAATGAAATTAATGAAAGAGATGTTAAATGACATTTCTTCTGAAGCGGAGA 1662
D 1604 TCGAAGAAATTAATGAAATTAATGAAAGAGATGTTAAATGACATTTCTTCTGAAGCGGAGA 1663
D 1605 TCGAAGAAATTAATGAAATTAATGAAAGAGATGTTAAATGACATTTCTTCTGAAGCGGAGA 1664
QY 477 TCCAAAGAAATGCTTTGACCTTTTACGCGCTGAAACAAAGTGGACACCAATATGCAATGAAC 418
D 478 TCCAAAGAAATGCTTTGACCTTTTACGCGCTGAAACAAAGTGGACACCAATATGCAATGAAC 419
D 479 TCCAAAGAAATGCTTTGACCTTTTACGCGCTGAAACAAAGTGGACACCAATATGCAATGAAC 420
QY 1663 CCGTGCATATTAATGATGAATTCGCTCGCGAAATFACACTCAATTAATAAATTAATAATTCCTG 1722
D 1664 CCGTGCATATTAATGATGAATTCGCTCGCGAAATFACACTCAATTAATAAATTAATAATTCCTG 1723
D 1665 CCGTGCATATTAATGATGAATTCGCTCGCGAAATFACACTCAATTAATAAATTAATAATTCCTG 1724
QY 417 GAACAGATGCTAAACAAGTAGTCTACCAAGTATTTATTAATCAATTCGTTGTTATCTG 358
D 418 GAACAGATGCTAAACAAGTAGTCTACCAAGTATTTATTAATCAATTCGTTGTTATCTG 359
D 419 GAACAGATGCTAAACAAGTAGTCTACCAAGTATTTATTAATCAATTCGTTGTTATCTG 360
QY 1723 ATTATCAATTTGATACCGAT---CTGACATCTATAAATCTAGAAAATTAATAATGAACCTC 1779
D 1724 ATTATCAATTTGATACCGAT---CTGACATCTATAAATCTAGAAAATTAATAATGAACCTC 1780
D 1725 ATTATCAATTTGATACCGAT---CTGACATCTATAAATCTAGAAAATTAATAATGAACCTC 1781
QY 357 AGTTTCATATCAGATGAGCAGATTTACCAACCGAAGTTGAGCTGAAATCTATCGTA 298
D 358 AGTTTCATATCAGATGAGCAGATTTACCAACCGAAGTTGAGCTGAAATCTATCGTA 299
D 359 AGTTTCATATCAGATGAGCAGATTTACCAACCGAAGTTGAGCTGAAATCTATCGTA 300
QY 1780 ATCAATTCGAACATTAATTTTAAATCTTATGTAATGCAATCAATTAATGATGTGAGA 1839
D 1781 ATCAATTCGAACATTAATTTTAAATCTTATGTAATGCAATCAATTAATGATGTGAGA 1840
D 1782 ATCAATTCGAACATTAATTTTAAATCTTATGTAATGCAATCAATTAATGATGTGAGA 1841
QY 297 ATCATTGGAACAATTCGTAATTTCTTCTTAGATAATGCGATTAAATTAATTCACCCGATC 238
D 298 ATCATTGGAACAATTCGTAATTTCTTCTTAGATAATGCGATTAAATTAATTCACCCGATC 239
D 299 ATCATTGGAACAATTCGTAATTTCTTCTTAGATAATGCGATTAAATTAATTCACCCGATC 240
```


Db 534557 TTCAAGCGGCTTAGCTTATCCATTTTCACAAAAAATAATTTACGTTACATCATCGTGGAAT 534616
QY 2048 TAAATTTAAAGTGAAATTAATAAGGAAACAACGCTTTAAATAACA 2091
Db 534617 CACTGTTGAAGTAAATCTGATAAAGGAAACAATTCAAAAGTCA 534660

RESULT 12
US-10-398-221-2058/c
; Sequence 2058, Application US/10398221
; Publication No. US200400185141
; GENERAL INFORMATION:
; APPLICANT: KUNST, Frederik
; APPLICANT: GLASER, Philippe
; TITLE OF INVENTION: Listeria innocua, genome and applications
; FILE REFERENCE: 344 702 - US
; CURRENT APPLICATION NUMBER: US/10/398,221
; CURRENT FILING DATE: 2003-03-27
; PRIOR APPLICATION NUMBER: PCT/FR 01/03 061
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: FR 00/12 697
; PRIOR FILING DATE: 2000-10-04
; NUMBER OF SEQ ID NOS: 4025
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 2058
; LENGTH: 3011208
; TYPE: DNA
; ORGANISM: Listeria innocua
US-10-398-221-2058

Query Match 7.0%; Score 154.8; DB 16; Length 3011208;
Best Local Similarity 44.5%; Pred. No. 3.9e-16;
Matches 927; Conservative 0; Mismatches 1112; Indels 45; Gaps 6;

QY 47 ATATTACCTTATTAGAAAAAGTCGTAATATAGGTGTACAAATGACGCCAAATTTTATAG 106
Db 2748229 ATATGATGCGATAAATAGGACGCGTAAGGTGAATAGAGAAATGAAACAAATTTTAGTAG 2748170

QY 107 TAGAAGATGAAACAAACTTAGCAAGATTTCTTGAATTTGGAACTCACACATGAAATATACA 166
Db 2748169 TAGATGATGACCGCAATTTCTTAACTTTAGGGCACTATCTCCGGCGAAGAGTTTTC 2748110

QY 167 ATGTGGACACAGATGATGATGGAACAAGCGTTTATAGATAAGCGCTTAGCCATTTACTATG 226
Db 2748109 ATGTGTTAGAAGCTAGTGTGCGGTAGAACGAGAAAGATAGTGGAAACGAGCAAGTTC 2748050

QY 227 ATTTAATCATATTAGATTTAATGTTGCGGTCAATTAATAGGCTTAGAAATTTGTCGCAAAA 286
Db 2748049 ATTTGGCGGTGATGTGATGATGCCCAACATGGATGGCTTTGAACATATGCAAAAAAAC 2747990

QY 287 T---TAGACAACAACATCTACACCTATCATTTATTAATTACAGCGAAAGTGATACGTATG 343
Db 2747989 TGGCATTAAGCTATCTCGATATCTCCCGTTATATGCTTGACAGCAAAAGATGCACTTGGCG 2747930

QY 344 ACAAGTTGCTGGCGCTTGATTAACGCTGACAGCATTTATATAGTTAAGCCGTTTGATATG 403
Db 2747929 ATAAATCGCGCGGTTTGAAGTGGGACGAGTACATATGCTAATACCTTTTGAACCCG 2747870

QY 404 AAGAACTTTTAGCAAGAAATTCGTGCAATTTTACGTCG-----TCAGCCCAAAAGATA 457
Db 2747869 AGGAGTTGATTTTTTCGAATACGGGCTTTTACTTAGACGTTCCCAATCAGGCTTAGTGAACGGA 2747810

QY 458 TTATCGATGTCACGGTATTACAAATTCATTAAGAACGCTTTTAAAGTCACCGTAATGGCG 517
Db 2747809 AAATTAATCTCGGAATATCACCATCGCAACCAAAAGATTAAGGATTAATAATTTGGCAAC 2747750

QY 518 CAGAAATTTGAATTAACAAAAACAGAGTATGATTTACTATATCTTCTAGCTGAAAAATAAAA 577
Db 2747749 AAGAACGGATGATTCCTGTAAAGAAATTTGAACGTTTATATCAATTAGCGAGCTATCCCG 2747690

QY 578 ACCATGTTATGCAACGGGAACAAATTTTAAATCATGATGCGGTTTATATAGTGAAGTAG 637
Db 2747689 GTCGGATTTTTTACTCGGGAAGAAATTAATTTGAACGGATTTTGGCAACGCGATTACGACGGA 2747630

QY 638 AAACAAATGTCGTAGATGTTTATATAAGATATTTACGAAACAGTTTAAACACCATAGCATC 697
Db 2747629 GTGACCGAACAGTAGATGTACATATTAACAGTTTGCAGCCCATTTTGTAGAGAAAGG 2747570

QY 698 GTGACAAAATGATTTGAAACAGTTTCGTGGCGTTGGGTATGTGTATACGATGACAAAACGTAA 757
Db 2747569 ACGGATTCGGATTTGTCAAGTTTCGGGTGTGGCTATATAAATTTGGAGGAAACACATGAA 2747510

QY 758 ATTCGCCAATAACTCGATTTATTTACCAACGATGATTACGTTTGTTCACGATATTTTGT 817
Db 2747509 GTCATTTATATAGTCGAATAGTTGTGACGATGTAGTTGTTCATTTCTTCCAGTAGCCCTGCT 2747450

QY 818 TTGTTTAAATTTATTT---ATTTTCTTTGAAAGATACACATGATGATGAGTGTGAGTGTGATGA 874
Db 2747449 CGGTTTTTCTTTGCGAACATTTTATCCAGATATAAATTTAAACCTTTTAAATGATGAAA 2747390

QY 875 TGCAGAACGAAGCTCAAGCGATATTAATAATTTTATTTTCAATTTCAA-----919
Db 2747389 AACGGCTAAGATAGCTGAGGAAAGTCCACAAATTTTATCAGTCGATAGGCGGATTTCTCT 2747330

QY 920 -----GCCTGTTTAAAGATATATCTGCATTTAGACTTTGAATGCACTTTTAGGTAA 967
Db 2747329 AAAAGAGTACTTTGGAATAATGTAGTGAACTTGGCTATGAACTGTATCTCAGCGATGCGAA 2747270

QY 968 TTTTCAAGAGATAATTTATTTATGATGAGCATATAATAAATTTATTTGAGACATCGAATGA 1027
Db 2747269 AGATCAATCGAGTTATTTCCGGTGGTGGCTTTGCAAAAAAAGATTTTACCTAACAAAAACGT 2747210

QY 1028 TAACAGATGAGAGTTTGAACACGAGTTATGAACACCGTTATTTTTCGCCGTATATAAAAA 1087
Db 2747209 TCAACAAGTTTGTAGCGGCGGAGACTTTATCAGCGGATTTGATACGTTTGTATACGCGCATTTT 2747150

QY 1088 ACCTATAAGGCAATTTGAATTTTAAATTTTAAAGAACCAATTTACAAACGACAGATTTTCAA 1147
Db 2747149 TATCAGAGATTTTGTGATAATGATGTCAGGAATACAAATTTGGTGTTCAGATTAAAGCGGA 2747090

QY 1148 AGGATAGTCTGTTTAAATTTCACTACAGAAATTTATGATTAACATCGTAAAAATCATTTGTA 1207
Db 2747089 CGGGAGCAACTGCGGCTTTTATAGTCAGAGTCCGGAACGCAATTCGGTGAACCTTAG 2747030

QY 1208 TATCATTCGCTGGCAATTTGGAGTGTATGCAACAATTTAATTAACCTGCCAATCATGTTATGT 1267
Db 2747029 AATCTTTTTCGGGATGATACCTTTGTTTTCATCTATTTATAGTATCTTATTTTGTGCTTAT 2746970

QY 1268 ATTTTCAACACAAATTTACTTAAACCGCTTCTCAGTTTATCAATAAATGATGAGATTCG 1327
Db 2746969 TAGTGGCCGATACATAGTGAACCCAGTTGTGAAGCTTACGAATGTCGAAAAAATCCG 2746910

QY 1328 ACAGATGCTTTTCAAAATTAATTTGCAATTTAAATACAAATTTATGAAGAAATAGATAATTT 1387
Db 2746909 GAGA---GGAATATAGATGTGGCGCTCGAAGTCCGCGTAAAGACGAGATTTGTCAGCT 2746853

QY 1388 AGCAATACGTTTAAATGAGATGATGAGCCAAATTTGAAGATCATTTTAAATCAACAAAGACA 1447
Db 2746852 GCGGACTCGTTTGCCAAATGCGGAGCGAACTCGAAAAATCTGAAGACGAGCCGTCAGGA 2746793

QY 1448 ATTTGTTGAAGATCGCTCATGAAATTTAGCAACCATTTACAAATTTATTTCAAGGTCATTT 1507
Db 2746792 ATTTGTCGAAATGTTTCTCATGAACCTACCACTTACATCGATGCAAGG---TTT 2746736

QY 1508 AAATTTGATTTCAGCGATGCGGAAAAAAGAACCCAGCATTTATGAAGAAATCGTTAAATAT 1567
Db 2746735 CGCCAGATTTTAGTTTCAGGAACACTAACCGAANAAGAACAAAGAAATATTTAACTGT 2746676

QY 1568 TTCTATTGAAGAAATGAATCGTATCATAAAAATTTAGTCGAAGATTAATTTGAAATTTGACTAA 1627
Db 2746675 TTATCCGAGGAGACGACACGCGCTTTCTTCTTTTAACGAAAAACAGTTACTTACACTTGCCTC 2746616

QY 1628 AGGAGATGTAATGACATTTCTTCTGAAGCGGACGCGTATTAATTAATGATGAATTCG 1687
Db 2746615 ACTTGCCAGGAATCTGAGTTACGGAANAAGAACCCAGTGCATTTAGCAGAACAGTGGCG 2746556

RESULT 14

US-10-282-122A-9474
; Sequence 9474, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9474
; LENGTH: 1839
; TYPE: DNA
; ORGANISM: Bacillus anthracis
US-10-282-122A-9474

Query Match 6.28; Score 136.8; DB 13; Length 1839;
Best Local Similarity 49.99; Pred. No. 3.1e-14;
Matches 398; Conservative 0; Mismatches 397; Indels 3; Gaps 2;
QY 1301 TTTATCAATTAATAATGATTGAGATTCGACGAGATGGTTTCAAAATAAATGCAATTA 1360
DB 1008 TTTATTAGATTTTATGACGAGAAATGAGCCATATATTTTACGCGTAGTTTCTGTAT 1067
QY 1361 TACAAATTTATGAAGAAATAGATAATTTAGCAAAATACGTTTAAATGAGATGATGAGCCAAAT 1420
DB 1068 TCAAAAGACACAGGAAAGCAATGTTTAAATGCGGTATTATATGATGCTGAGCA 1127
QY 1421 TGAAGATCATTTAATCAACAAAGACAAATTTGTTGAAGATGCGTCACATGAATTAAGAAC 1480
DB 1128 GGAGCGGTATGAAGAGGGAACCGCGGAGTTTGTAGCAAAAGCTCTCATGAATTAAGAAC 1187
QY 1481 ACCATTCAAAATTTATCAAGTCATTTAAATTTGATTCACCGATGGGAAAAAGACCC 1540
DB 1188 GCAATTAACGACATGCGGAGCTACTTAGAGGCACCTTACAGCGGTGCGGCAAGATCC 1247
QY 1541 AGCAGTATTAGAAGATCGTTAAATATTTCTATTGAAGAAATGAATCGTATCAATAAAT 1600

DB 1248 GAATATTGCACCGCAATTTTAAACAGTTATACAGGAAGAAACAGAAAGAAATGATTAGACT 1307
QY 1601 AGTCGAAGAAATTAATTTGAATTTGACTTAAGAGAGAT-GTAATGACATTTCTTCTTGAGCGC 1659
DB 1308 TGFAAATGCGTTAATGCAACTATCTAAACTAGATAGTACAGAAACATCGTTTAAATGAAGA 1367
QY 1660 AGACCGGTGCATATTAATGATGAAATTCGGTCGGGAATACACTCATTAATAAACAATTTGCATC 1719
DB 1368 ATGGTTCGACTTTACTGACTTCTTTAATACATTTATGTCGTTTGAATGCTAAAGA 1427
QY 1720 CTGATTATCAATTTGATACGGATCTGACATCTAAAAATCTAGA--AATTAAAAATGAACC 1777
DB 1428 GCAAAACGTAAGTTTCAACAGCATCTTCTCTAAAAAATCTCGATTTATGATATGGATAC 1487
QY 1778 TCATCAATTCGACAAATTTTAAATCTTTATTTATGATTAATGCAATCAATATCATGTGAA 1837
DB 1488 GGATAAATTTACGCAAGTATTGTATAACATTTATTCGAATGCAATTAAGTATTCACCAGA 1547
QY 1838 GAATAAGAAATTAAGTTTAAGACAAGTTTAAAAAATAAGCAAAAAATTAATGAAATTAC 1897
DB 1548 AGTGAACAGTAACATATCTGTTTACGTGATCCGGAGAAATTAATAGAGATTAGTGAAG 1607
QY 1898 AGATCATGGAATTTGGTATTTCCAGAGGAAGATCAAGATTTTCATTTTTCATCGCTTTATCG 1957
DB 1608 TGACCAAGGAATGGGTATTTCCGAAAGAAACGTCGATAAATCTTTGAACGTTTATCG 1667
QY 1958 AGTGGATAATCTCGTTCAAGAAGTCAAGCGGTAAATGACATCGGATTCATTTGCTCA 2017
DB 1668 TGTAGATAAGCGCGCTCACGAAATGGGTGTGATACAGGACTTGGATTAAGCAATTCGAA 1727
QY 2018 AAAAATCATTAATTAACGGAGGATCGATTAAAAATTTAAAGTGAATTAATAAAGGAAC 2077
DB 1728 AGAAATGATTGAGGCACNTGGTCTCGATTGGGCGAGAGTGAAGAGAAAGGAC 1787
QY 2078 AACGTTTAAATCATATT 2095
DB 1788 AACGATTATTTCACATT 1805

RESULT 15

US-10-282-122A-35146
; Sequence 35146, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23

```
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 35146
; LENGTH: 1755
; TYPE: DNA
; ORGANISM: Staphylococcus haemolyticus
US-10-282-122A-35146

Query Match      6.1%; Score 133.4; DB 13; Length 1755;
Best Local Similarity 51.3%; Pred. No. 1.2e-13;
Matches 339; Conservative 0; Mismatches 316; Indels 6; Gaps 1;

Qy 1441 AAAGACAATTGTTGAAGATCGTCACATGAATTACGAACACCATTACAAATTATTCAAG 1500
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1079 AGAAGATTTCATGCCAATGTTTCACAIGAAATTACGTACACCTATTTCGTTATTACAAG 1138
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1501 GTCATTTAAATTGATTCCAGGATGGGGAAGAAAAGACCCAGCAGTATTAGAAGAAATCGT 1560
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1139 GTTACACTGAATCAATTGTAGATGGTATTGTCACTGAACACAGAGAAATTCGAGATTCCT 1198
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1561 TAAATATTTCTATTGAAGAAATGAATCGTATCATAAATATTAGTCGAAGAAATCTTGAAT 1620
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1199 TATCAATTTGTATTAGATGAAACGAAACCGTTTAAATCGATTAGTAAATGAATTAATAATG 1258
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1621 TGACTAAAGGAGATGTAATGACATTTCTTCTGAAGCCGACACCGTCATATTAATGATG 1680
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1259 TGGTCGCATCGATGCAGAGGTTTAACTGTTAATAAGAAATAACAACTATTAAACCTT 1318
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1681 AAATTCGCTCGCAATACACTCATTTAAACAAATTCCTCTGATTTATCAATTTGATACGG 1740
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1319 TACTATCTAAAAATGCAAAATGAAATATAGACAACAGGCTGAGGATTTAGAAATGAATATGT 1378
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1741 ATCTGACATCTAAAA-----ATCTAGAAATTAATAAGAAACCTCATCAATTCGAACAAAT 1794
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1379 CATTAGAACCAAAACATCGATGATGAATTTGTGGGAATACGATCCGATCGCATAGACCAAG 1438
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1795 TATTTTAAATCTTTTTCATATGCAATCAATATATGATGTGAAGAAATAAGAAATTAAG 1854
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1439 TTCTAACAAATTTAATTGACAAATGCCACACGTTTACTCAACCCAGGTGATTCAATATCAA 1498
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1855 TTAAGACAAGGTTTAAAAAATAAGCAAAAAATAATGAAATTTACAGATCATGGAATTTGGTA 1914
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1499 TTACAACAACCTACGGATGATTCCTATCAACGTTGTACATTAAAGATACCTGTTAGTGCA 1558
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1915 TTCAGAGGAAGATCAAGATTTCTTTTGTATCGCTTTTATCGAGTGGAATAAATCTCGTT 1974
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1559 TATCTCCAGAGACTCTTGAACGCTATTTCGACAGATTTTATAAAGTAGAGGCTTCAAGAA 1618
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1975 CAAGAAGTCAGAGCGGTAATGATCGGATTTCTATTGCTCAAAAATCATTCAATTA 2034
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1619 CTAGAGGTAAACAGGGAACTGGTTTAGGACTATTATTCTGTAAAATGATAATCGAAGAAC 1678
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 2035 ACGGAGGATCGATTAAATAATAAAGTGAATTAATAAAGGAACCAACGTTTAAATCATAT 2094
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1679 ATGGTGGTACTATTAAAGTGAAGTAAAGTAAACAAAGGTACAACATTCAATTAATAGT 1738
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 2095 T 2095
Db |
Qy 1739 T 1739
Db |
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 5, 2004, 00:16:04 ; Search time 8818.24 Seconds
(without alignments)
10818.262 Million cell updates/sec

Title: US-09-006-627-1
Perfect score: 2201
Sequence: 1 TAAATTTAAAGCAACTATT.....AAAACAATAACAGTGGTT 2201

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671515995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

1: gb.ba.*

2: gb.htg.*

3: gb.in.*

4: gb.om.*

5: gb.ov.*

6: gb.pat.*

7: gb.ph.*

8: gb.pl.*

9: gb.pr.*

10: gb.ro.*

11: gb.sts.*

12: gb.sy.*

13: gb.un.*

14: gb.vi.*

15: em.ba.*

16: em.fun.*

17: em.hum.*

18: em.in.*

19: em.mu.*

20: em.om.*

21: em.or.*

22: em.ov.*

23: em.pat.*

24: em.ph.*

25: em.pl.*

26: em.ro.*

27: em.sts.*

28: em.un.*

29: em.vi.*

30: em.htg.hum.*

31: em.htg.inv.*

32: em.htg.other.*

33: em.htg.mus.*

34: em.htg.pln.*

35: em.htg.rod.*

36: em.htg.mam.*

37: em.htg.vrt.*

38: em.sy.*

39: em.htgo.hum.*

40: em.htgo.mus.*

41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2201	100.0	2201	6	AR306346	Sequence
2	2201	100.0	2201	6	BD136060	Histidine
3	2201	100.0	346900	1	AP003362	Staphyloc
4	2156.2	99.8	3083	1	AF165314	Staphyloc
5	2183.4	99.2	295350	1	AP004826	Staphyloc
6	2159	98.1	303750	1	AP003133	Staphyloc
7	1432	65.1	1440	6	AR354607	Sequence
8	1349.8	61.3	1353	6	AX621580	Sequence
9	1108.2	50.3	300892	1	AE016747	Staphyloc
10	1085	49.3	3230	1	AF269630	Staphyloc
11	1085	49.3	3230	6	AX144950	Sequence
12	1041.6	47.3	3018	1	AF270244	Staphyloc
13	1041.6	47.3	3018	6	AX145562	Sequence
14	978.8	44.5	3099	1	AF269746	Staphyloc
15	978.8	44.5	3099	6	AX145064	Sequence
16	736	33.4	736	6	AE3826	Sequence
17	736	33.4	736	6	AR306347	Sequence
18	736	33.4	736	6	BD136061	Histidine
19	673.2	30.6	4858	6	AR354261	Sequence
20	657	29.9	657	6	AX621578	Sequence
21	471	21.4	855	6	AX142703	Sequence
22	471	21.4	855	6	AX144067	Sequence
23	399.8	18.2	195269	6	AX417035	Sequence
24	399.8	18.2	333050	1	AL596168	Listeria
25	399.8	18.2	349980	6	AX413017	Sequence
26	399.8	18.2	349980	6	AX417044	Sequence
27	389.8	16.8	250050	1	AL591978	Listeria
28	389.8	16.8	349980	6	AX641668	Sequence
29	355.4	16.1	3619	1	AF139908	Listeria
30	307	13.9	11726	1	AE010631	Sequence
31	302.6	13.7	501	6	AX142701	Sequence
32	302.6	13.7	501	6	AX144065	Sequence
33	289.8	13.2	10816	1	AE013066	Thermoana
34	285.2	13.0	293264	1	AE017031	Bacillus
35	269.2	12.2	301488	1	AE016950	Enterococ
36	262.8	11.9	10894	1	AE010570	Fusobacte
37	234.8	10.7	303250	1	AP003193	Clostridi
38	232.2	10.5	5001	1	LLU81166	U81166 Lactococcus
39	223.8	10.2	10575	1	AE006390	Lactococc
40	218.6	9.9	300050	1	AP004595	Oceanobac
41	215.2	9.8	290433	1	AE017034	Bacillus
42	212.8	9.7	20601	1	AE014265	Streptoco
43	212.4	9.7	44145	6	AX602195	Sequence
44	212.4	9.7	174050	1	SAG766852	Streptoco
45	209.2	9.5	57272	6	AX602202	Sequence

ALIGNMENTS

RESULT 1
AR306346
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL

AR306346
Sequence 1 from patent US 6548281.
AR306346
AR306346.1 GI:31696114
Unknown.
Unclassified.
1 (bases 1 to 2201)
O'Dwyer,K.M., Jaworski,D.D., Mooney,J.L., Shilling,L.K.,
Throup,J.P., Wallis,N.G., Zhong,Y.Y. and Wang,M.
Histidine kinase
Patent: US 6548281-A 1 15-APR-2003;

2201 bp
DNA
linear
PAT 12-JUN-2003

FEATURES		Location/Qualifiers	
source	1..2201	/organism="unknown"	
		/mol_type="genomic DNA"	
ORIGIN			
Query Match 100.0%; Score 2201; DB 6; Length 2201;			
Best Local Similarity 100.0%; Pred. No. 2.1e-307;			
Matches 2201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	1	TAATTTAAAGCAACTATTGTATAGAAAATACAAAATTTAAATATATATACCTATTATTA	60
Db	1	TAATTTAAAGCAACTATTGTATAGAAAATACAAAATTTAAATATATATATACCTATTATTA	60
Qy	61	GAAGAGTCGTATATAGAGGTGTACAAATGACGCAAAATTTTAAATAGTAGAAGATGAACAA	120
Db	61	GAAGAGTCGTATATAGAGGTGTACAAATGACGCAAAATTTTAAATAGTAGAAGATGAACAA	120
Qy	121	AACCTAGCAAGATTTCTTGAATTTGGAATCTCACATGAAAATTTACAAATGTGGACACAGAG	180
Db	121	AACCTAGCAAGATTTCTTGAATTTGGAATCTCACATGAAAATTTACAAATGTGGACACAGAG	180
Qy	181	TATGATGGACACAGCGTTTACATAAAGCGCTTAGCCATTACTATGATTTAAATCATATTA	240
Db	181	TATGATGGACACAGCGTTTACATAAAGCGCTTAGCCATTACTATGATTTAAATCATATTA	240
Qy	241	GATTTAATGTTGCGGTCAAATTAATGCGCTTAGAAAATTTGTCGCAAAATTTAGACAAACAA	300
Db	241	GATTTAATGTTGCGGTCAAATTAATGCGCTTAGAAAATTTGTCGCAAAATTTAGACAAACAA	300
Qy	301	TCTACACCTATCATTAATTAACAGCGAAAGTGATACGATGACAAAGTTGCTGGGCTT	360
Db	301	TCTACACCTATCATTAATTAACAGCGAAAGTGATACGATGACAAAGTTGCTGGGCTT	360
Qy	361	GATTACGGTCGACAGCAATTTATAGTTAAGCGCTTGATATGAAAGAACTTTTAGCAAGA	420
Db	361	GATTACGGTCGACAGCAATTTATAGTTAAGCGCTTGATATGAAAGAACTTTTAGCAAGA	420
Qy	421	ATTGCGTCAATTTTACGTCGACGCAAAAGGATATTATCGATGTCGCAACGGTATTACA	480
Db	421	ATTGCGTCAATTTTACGTCGACGCAAAAGGATATTATCGATGTCGCAACGGTATTACA	480
Qy	481	ATTGATAAGAACGCTTTTAAAGTGACGGTAAATGGCGCAGAAATTTGAATTAACAAACAA	540
Db	481	ATTGATAAGAACGCTTTTAAAGTGACGGTAAATGGCGCAGAAATTTGAATTAACAAACAA	540
Qy	541	GAGTATGATTTACTATATCTTTAGCTGAAAATAAAAAACCATGTTATGCAACGGGAACAA	600
Db	541	GAGTATGATTTACTATATCTTTAGCTGAAAATAAAAAACCATGTTATGCAACGGGAACAA	600
Qy	601	ATTTTAAATCATGTATGGGGTTATATAGTGAAGTAGAAGCAAAATGTCGTAGATGTTTAT	660
Db	601	ATTTTAAATCATGTATGGGGTTATATAGTGAAGTAGAAGCAAAATGTCGTAGATGTTTAT	660
Qy	661	ATAAGATATTACGAAACAGTTTAAACCATACGATCGTCGCAAAATTTGAATGAACAGATT	720
Db	661	ATAAGATATTACGAAACAGTTTAAACCATACGATCGTCGCAAAATTTGAATGAACAGATT	720
Qy	721	CGTGGCGTTGGGTATGTGATACGATGACAAAACGTTAAATTTGCGCAATTAATGATTTTG	780
Db	721	CGTGGCGTTGGGTATGTGATACGATGACAAAACGTTAAATTTGCGCAATTAATGATTTTG	780
Qy	781	TTACACGATGATTAACGTTTGTACGATATTTTGTGTTTGTGTTTAAATTTATTTTCT	840
Db	781	TTACACGATGATTAACGTTTGTACGATATTTTGTGTTTGTGTTTAAATTTATTTTCT	840
Qy	841	TGAAGATACACTGCATTAATAGTACGCTTCATGATGACGAAACGAAAGCTCAAGCGATTA	900
Db	841	TGAAGATACACTGCATTAATAGTACGCTTCATGATGACGAAACGAAAGCTCAAGCGATTA	900
Qy	901	ATAATTTATTTCAATTCGAGCTGTTAAAGATATATCTGCAATTTAGACTTGAATGCATCTT	960
Db	901	ATAATTTATTTCAATTCGAGCTGTTAAAGATATATCTGCAATTTAGACTTGAATGCATCTT	960
Qy	961	TAGGTAATTTTCAAGAGATAATTTATTTATGATGAGCATATATATATTTTGGACAT	1020
Db	961	TAGGTAATTTTCAAGAGATAATTTATTTATGATGAGCATATATATATTTTGGACAT	1020
Qy	1021	CGAATGATAACACAGTGAGAGTTGAACACAGGTTTATGAACACCGTTATTTTGGACGGTAA	1080
Db	1021	CGAATGATAACACAGTGAGAGTTGAACACAGGTTTATGAACACCGTTATTTTGGACGGTAA	1080
Qy	1081	TAAAAAACGCTATATAAGGCATTGAATATTTAATTTAAAGAACCAATTTCAACGCAAG	1140
Db	1081	TAAAAAACGCTATATAAGGCATTGAATATTTAATTTAAAGAACCAATTTCAACGCAAG	1140
Qy	1141	ATTTCAAGGGTATAGCTTTGTTAAATTTCACTAGAAAATTTATGATAAATCGTAAAAAT	1200
Db	1141	ATTTCAAGGGTATAGCTTTGTTAAATTTCACTAGAAAATTTATGATAAATCGTAAAAAT	1200
Qy	1201	CATTGTATATCATTTGGCGCTGGCAATTTGGAGTGATTCGAAACAAATTAATACTGCCACAATCA	1260
Db	1201	CATTGTATATCATTTGGCGCTGGCAATTTGGAGTGATTCGAAACAAATTAATACTGCCACAATCA	1260
Qy	1261	GTTATGTATTTTCAACACAAATTTACTAAACCGCTTGTCAAGTTTATCAAAATAAATGATTG	1320
Db	1261	GTTATGTATTTTCAACACAAATTTACTAAACCGCTTGTCAAGTTTATCAAAATAAATGATTG	1320
Qy	1321	AGATTCGACGAGATGCTTTTCAAAATAAATTTGCAATTTAAATACAAATTTATGAAGAAATAG	1380
Db	1321	AGATTCGACGAGATGCTTTTCAAAATAAATTTGCAATTTAAATACAAATTTATGAAGAAATAG	1380
Qy	1381	ATAATTTAGCAATACGTTTAAATGAGATGATGAGCCAAATTTGAAGAAATCATTTAATCAAC	1440
Db	1381	ATAATTTAGCAATACGTTTAAATGAGATGATGAGCCAAATTTGAAGAAATCATTTAATCAAC	1440
Qy	1441	AAAGACAATTTTGTGTAAGATGCGTCACATGAAATTTACGAAACCAATTTTATTTCAAG	1500
Db	1441	AAAGACAATTTTGTGTAAGATGCGTCACATGAAATTTACGAAACCAATTTTATTTCAAG	1500
Qy	1501	GTCAATTTAAATTTGATTTACGCGATGGGAAACCAACCGCAGTATTAGAGAATCGT	1560
Db	1501	GTCAATTTAAATTTGATTTACGCGATGGGAAACCAACCGCAGTATTAGAGAATCGT	1560
Qy	1561	TAAATTTCTATTGAAGAAATGAATCGTATCATATAAATTTAGTCGAAGAAATTAATTTGAAT	1620
Db	1561	TAAATTTCTATTGAAGAAATGAATCGTATCATATAAATTTAGTCGAAGAAATTAATTTGAAT	1620
Qy	1621	TGACTAAAGGAGATGTAATGACATTTCTTCTGAAGCGCAGACCGTGCATATTAATGATG	1680
Db	1621	TGACTAAAGGAGATGTAATGACATTTCTTCTGAAGCGCAGACCGTGCATATTAATGATG	1680
Qy	1681	AAATTCGCTCGGAATACACTCATTAACCAATTTGCAATCTGATTTATCAATTTGATACGG	1740
Db	1681	AAATTCGCTCGGAATACACTCATTAACCAATTTGCAATCTGATTTATCAATTTGATACGG	1740
Qy	1741	ATCTGACATCTAAAAATCTAGAAAATTTAAATGAAACCTCATCAATTCGAAATTTATTTT	1800
Db	1741	ATCTGACATCTAAAAATCTAGAAAATTTAAATGAAACCTCATCAATTCGAAATTTATTTT	1800
Qy	1801	TAACTCTTTATGATAATGCAATCAAAATATGATGTGAAGAAATAAGAAAAATTTAAAGTTAAGA	1860
Db	1801	TAACTCTTTATGATAATGCAATCAAAATATGATGTGAAGAAATAAGAAAAATTTAAAGTTAAGA	1860
Qy	1861	CAGGTTTAAATTAAGCAAAATTAATTTGAATTTACAGATCATGAAATTTGGTATTTCCAG	1920
Db	1861	CAGGTTTAAATTAAGCAAAATTAATTTGAATTTACAGATCATGAAATTTGGTATTTCCAG	1920
Qy	1921	AGGAAGATCAAGATTTTCAATTTTGTATCGCTTTTATCGAGTGAATAAATCTCGTTCAAGAA	1980
Db	1921	AGGAAGATCAAGATTTTCAATTTTGTATCGCTTTTATCGAGTGAATAAATCTCGTTCAAGAA	1980
Qy	1981	GTCAAGCGCGTAAATGACCTCGGATTTATTTGCTCAAAAAATCATTTCAATTTAAACGGAG	2040
Db	1981	GTCAAGCGCGTAAATGACCTCGGATTTATTTGCTCAAAAAATCATTTCAATTTAAACGGAG	2040

```
Qy 2041 GATCATTAAATTAAGAGTGAAATTAATAAGGAACACCGTTTAAATATCATATTTTAAAT 2100
Db 2041 GATCATTAAATTAAGAGTGAAATTAATAAGGAACACCGTTTAAATATCATATTTTAAAT 2100

Qy 2101 CATGCTCAGACGCGCAATCAAGTCATAGGATCAATTTTAAAGTACACATTAGCTGTGA 2160
Db 2101 CATGCTCAGACGCGCAATCAAGTCATAGGATCAATTTTAAAGTACACATTAGCTGTGA 2160

Qy 2161 CTAATGTATAAGAAACAACTATATAAAACAAATAAACAAGTGGTT 2201
Db 2161 CTAATGTATAAGAAACAACTATATAAAACAAATAAACAAGTGGTT 2201

RESULT 2
BD136060
LOCUS BD136060 2201 bp DNA linear PAT 18-SEP-2002
DEFINITION Histidine kinase.
ACCESSION BD136060
VERSION BD136060.1 GI:23231005
KEYWORDS JP 2002508955-A/1.
SOURCE Staphylococcus aureus
ORGANISM Staphylococcus aureus
Bacteria; Firmicutes; Bacillales; Staphylococcus.
REFERENCE 1 (bases 1 to 2201)
AUTHORS Wallis,N.G., Schilling,L.K., Mooney,J.L., Debouck,C., Twon,I.I.,
Jaworski,D.D., Wang,M. and Throup,J.P.
TITLE Histidine kinase
JOURNAL Patent: JP 2002508955-A 1 26-MAR-2002;
COMMENT SMITHKLINE BEECHAM CORP
OS Staphylococcus aureus
PN JP 2002508955-A/1
PD 26-MAR-2002
PF 12-JAN-1999 JP 2000540212
PR 13-JAN-1998 US 09/006627
PI NICOLA G WALLIS, LISA K SCHILLING, JEFFREY L MOONEY, CHRISTINE
PI DEBOUCK,
PI I I TWON, DEBORAH D JAWORSKI, MING WANG, JOHN P THROUP PC
C12N15/09,A61K38/00,A61K38/51,A61K39/02,A61K39/395,A61K39/395, PC
A61K45/00,
PC A61P31/04,A61P37/04,C07H21/02,C07K16/40,C12N1/21,C12N5/10, PC
C12N9/12,
PC C12Q1/48,C12Q1/68,C12N15/00,A61K37/02,A61K37/56,C12N5/00 CC
Histidine kinase
FH Key Location/Qualifiers
FT source 1..2201
FT /organism='Staphylococcus aureus'.
FEATURES
source
Location/Qualifiers
1..2201
/organism='Staphylococcus aureus'
/mol_type='genomic DNA'
/db_xref='taxon:1280'

ORIGIN
Query Match 100.0%; Score 2201; DB 6; Length 2201;
Beat Local Similarity 100.0%; Pred. No. 2.1e-307;
Matches 2201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAAATTTAAAGACCACTATGATAGAAAATACAAAATTTTAAATATATTTACCTTATTA 60
Db 1 TAAATTTAAAGACCACTATGATAGAAAATACAAAATTTTAAATATATTTACCTTATTA 60

Qy 61 GAAAAGTCGTAATATAGAGTGATACAAATGACGCAAAATTTTAAATAGTAGAAGATGAACAA 120
Db 61 GAAAAGTCGTAATATAGAGTGATACAAATGACGCAAAATTTTAAATAGTAGAAGATGAACAA 120

Qy 121 AACTTAGCAAGATTTCTTTGAATTGGAATCTACACATGAAAATTTACAAATGTGGACACAGAG 180
Db 121 AACTTAGCAAGATTTCTTTGAATTGGAATCTACACATGAAAATTTACAAATGTGGACACAGAG 180

Qy 181 TATGATGCAAGACGGTTTATAGATAAAGCGCTTAGCCATTACTATGATTTTATCATATTA 240
Db 181 TATGATGCAAGACGGTTTATAGATAAAGCGCTTAGCCATTACTATGATTTTATCATATTA 240
```

```
Qy 241 GATTTAATGTGGCGTCAATTAATTAATGGCTTAGAAATTTGTCGCAAAATTTAGACAACAACAA 300
Db 241 GATTTAATGTGGCGTCAATTAATTAATGGCTTAGAAATTTGTCGCAAAATTTAGACAACAACAA 300

Qy 301 TCTACACCTATCATTATAATTAAGCGAAAAGTGATACGTTAGCAAAAGTTGCTGGGCTT 360
Db 301 TCTACACCTATCATTATAATTAAGCGAAAAGTGATACGTTAGCAAAAGTTGCTGGGCTT 360

Qy 361 GATTACCGGTGACAGCGATTATATAGTTTAAGCGCTTTCGATATTGAAGAACTTTTAGCAAGA 420
Db 361 GATTACCGGTGACAGCGATTATATAGTTTAAGCGCTTTCGATATTGAAGAACTTTTAGCAAGA 420

Qy 421 ATTCTGCAATTTTACGTCGACGCCACAAAAGGATATTATCGATGTCAACGGTATTACA 480
Db 421 ATTCTGCAATTTTACGTCGACGCCACAAAAGGATATTATCGATGTCAACGGTATTACA 480

Qy 481 ATTGATAAGAACCGCTTTTAAAGTGACGGTAAATGGCGCAGAAATTTGAATTAACAAAACA 540
Db 481 ATTGATAAGAACCGCTTTTAAAGTGACGGTAAATGGCGCAGAAATTTGAATTAACAAAACA 540

Qy 541 GAGTATGATTACTATATCTTCTAGCTGAAAAATAAAAAACCATGTTATGCAACGGGAACAA 600
Db 541 GAGTATGATTACTATATCTTCTAGCTGAAAAATAAAAAACCATGTTATGCAACGGGAACAA 600

Qy 601 ATTTTAAATCATGTATCGGGTTATAATAGTGAAGTAGAAAACAAATGTCGTAGATGTTTAT 660
Db 601 ATTTTAAATCATGTATCGGGTTATAATAGTGAAGTAGAAAACAAATGTCGTAGATGTTTAT 660

Qy 661 ATAAGATATTACGAAACAAGTTTAAACCATACGATCGTGACAAAATGATTGAACAGTT 720
Db 661 ATAAGATATTACGAAACAAGTTTAAACCATACGATCGTGACAAAATGATTGAACAGTT 720

Qy 721 CGTGGCGTTGGGTATGTATACGATGACAAAACGTAATTTGCGCAATAAATGCGGATTTTG 780
Db 721 CGTGGCGTTGGGTATGTATACGATGACAAAACGTAATTTGCGGCAATAAATGCGGATTTTG 780

Qy 781 TTACCACGATGATTAACGTTTGTGTCACGATATTTTGTGTTTGTATTAATTTATTTTCT 840
Db 781 TTACCACGATGATTAACGTTTGTGTCACGATATTTTGTGTTTGTATTAATTTATTTTCT 840

Qy 841 TGAAGAATACACTGCATTAATAGTGAGCTTGATGATGACGAACGAGCTCAAGCCATATTA 900
Db 841 TGAAGAATACACTGCATTAATAGTGAGCTTGATGATGACGAACGAGCTCAAGCCATATTA 900

Qy 901 ATAATTTATTTTCAATTTAAGCTGTTAAAGATATATCTGCATTTAGACTTGAATGCATCTT 960
Db 901 ATAATTTATTTTCAATTTAAGCTGTTAAAGATATATCTGCATTTAGACTTGAATGCATCTT 960

Qy 961 TAGGTAAATTTTCAAGAGATAATTTATTTATGATGAGCATAATAATAATTTATTTGAGACAT 1020
Db 961 TAGGTAAATTTTCAAGAGATAATTTATTTATGATGAGCATAATAATAATTTATTTGAGACAT 1020

Qy 1021 CGAATGATAACAGTGAGAGTTGAACCGGTTATGAACACCGTTATTTTGACCGCGTAA 1080
Db 1021 CGAATGATAACAGTGAGAGTTGAACCGGTTATGAACACCGTTATTTTGACCGCGTAA 1080

Qy 1081 TAAAAAACCGCTATAAAGGCATTTGAATTTTAAATTTTAAAGAACCAATTTACAACGCAAG 1140
Db 1081 TAAAAAACCGCTATAAAGGCATTTGAATTTTAAATTTTAAAGAACCAATTTACAACGCAAG 1140

Qy 1141 ATTTCAAAGGATATAGCTTGTAAATTCATTCACCTAGAAAAATTTATGATAAATTCGTAAT 1200
Db 1141 ATTTCAAAGGATATAGCTTGTAAATTCATTCACCTAGAAAAATTTATGATAAATTCGTAAT 1200

Qy 1201 CATTGATATCATTTGCGCTGGCATTTGGAGTGATTGCAACAATTTAATCTGCCCAATCA 1260
Db 1201 CATTGATATCATTTGCGCTGGCATTTGGAGTGATTGCAACAATTTAATCTGCCCAATCA 1260

Qy 1261 GTTATGATTTTCAACACAAATTTACTAAACCGCTTGTACGTTTATCAAAATAAATGATTG 1320
Db 1261 GTTATGATTTTCAACACAAATTTACTAAACCGCTTGTACGTTTATCAAAATAAATGATTG 1320

Qy 1321 AGATTCCGACGAGATGGTTTCAAATAAATTTGCAATTTAAATTAACAAATTTATGAAGAAATAG 1380
```

|||||
1321 AGATTCACGAGATGGTGTTCCTCAAAATAAATGCAATTAATAACAATTTATGAAGAAATAG 1380
Qy |||||||
1381 ATAATTTAGCAATACGTTTAAATGAGATGATGAGCCAAATTTGAAGAAATCAATTTAATCAAC 1440
Db |||||||
1381 ATAATTTAGCAATACGTTTAAATGAGATGATGAGCCAAATTTGAAGAAATCAATTTAATCAAC 1440
Qy |||||||
1441 AAAGACAATTTGTTGAAGATGCGTCACATGAATTAAGAACACCAATTAACAATTTATCAAG 1500
Db |||||||
1441 AAAGACAATTTGTTGAAGATGCGTCACATGAATTAAGAACACCAATTAACAATTTATCAAG 1500
Qy |||||||
1501 GTCAATTTAAATTTGATTGACGATGGGCAAAAGAACCCAGCAGATTAAGAAAGATTCGT 1560
Db |||||||
1501 GTCAATTTAAATTTGATTGACGATGGGCAAAAGAACCCAGCAGATTAAGAAAGATTCGT 1560
Qy |||||||
1561 TAAATATTTCTATTGAAAGAAATGAATCGTATCAATAAATTTAGTCGAAGAAATTAATTTGAAT 1620
Db |||||||
1561 TAAATATTTCTATTGAAAGAAATGAATCGTATCAATAAATTTAGTCGAAGAAATTAATTTGAAT 1620
Qy |||||||
1621 TGACTAAAGAGATGTAATGACATTTCTTCTGAAGCGCAGACCGTCATATTAATGATG 1680
Db |||||||
1621 TGACTAAAGAGATGTAATGACATTTCTTCTGAAGCGCAGACCGTCATATTAATGATG 1680
Qy |||||||
1681 AAATTCGCTCGGAATACACTTAAATAAACAATTTGCAATCTGATTAATCAATTTGATACGG 1740
Db |||||||
1681 AAATTCGCTCGGAATACACTTAAATAAACAATTTGCAATCTGATTAATCAATTTGATACGG 1740
Qy |||||||
1741 ATCTGACATCTAAATAATCTAGAAATTAATAAATGAACCTCATCAATTCGAACAATTTATTTT 1800
Db |||||||
1741 ATCTGACATCTAAATAATCTAGAAATTAATAAATGAACCTCATCAATTCGAACAATTTATTTT 1800
Qy |||||||
1801 TAATCTTTTATGTAATGCAATCAATATGATGTAAGAAATTAAGAAATTTAAGATTAAGA 1860
Db |||||||
1801 TAATCTTTTATGTAATGCAATCAATATGATGTAAGAAATTAAGAAATTTAAGATTAAGA 1860
Qy |||||||
1861 CAAGTTTAAATAATTAAGCAAAATAATTAAGAAATTAAGAAATTAAGAAATTTAAGATTTCCAG 1920
Db |||||||
1861 CAAGTTTAAATAATTAAGCAAAATAATTAAGAAATTAAGAAATTAAGAAATTTAAGATTTCCAG 1920
Qy |||||||
1921 AGGAAGATCAAGATTTTCAATTTTGTATCGCTTTTATGAGTGGATTAATCTCGTTCAAGAA 1980
Db |||||||
1921 AGGAAGATCAAGATTTTCAATTTTGTATCGCTTTTATGAGTGGATTAATCTCGTTCAAGAA 1980
Qy |||||||
1981 GTCAAGCGGTATGAGTCTGATTAATCTGCTCAAAATAATCAATTAAGCGGAG 2040
Db |||||||
1981 GTCAAGCGGTATGAGTCTGATTAATCTGCTCAAAATAATCAATTAAGCGGAG 2040
Qy |||||||
2041 GATCGATTAATAATTAAGTGAATTAATAAGGAACAACGTTTAAATCATATTTTAAAT 2100
Db |||||||
2041 GATCGATTAATAATTAAGTGAATTAATAAGGAACAACGTTTAAATCATATTTTAAAT 2100
Qy |||||||
2101 CATGCTGTGAGACGTCATCAATCAAGTCATAGGATCAATTTTAAAGTACACATTAAGTGTGA 2160
Db |||||||
2101 CATGCTGTGAGACGTCATCAATCAAGTCATAGGATCAATTTTAAAGTACACATTAAGTGTGA 2160
Qy |||||||
2161 CTAATGTAAGAACACTATAAACAATAAACAATTAACAGTGGTT 2201
Db |||||||
2161 CTAATGTAAGAACACTATAAACAATAAACAATTAACAGTGGTT 2201

RESULT 3
AP003362/c 346900 bp DNA linear BCT 07-FEB-2002
LOCUS Staphylococcus aureus subsp. aureus Mu50 genomic DNA, complete
DEFINITION sequence, section 5/9.
ACCESSION AP003362 BA000017
VERSION AP003362.2 GI:14247083
KEYWORDS Staphylococcus aureus subsp. aureus Mu50
SOURCE Staphylococcus aureus subsp. aureus Mu50
ORGANISM Bacteria; Firmicutes; Bacillales; Staphylococcus.
REFERENCE 1
AUTHORS Kuroda,M., Ohta,T., Uchiyama,I., Baba,T., Yuzawa,H., Kobayashi,I.,

Cui L., Oguchi A., Aoki K., Nagai Y., Lian J., Ito T., Kanamori M.,
Matsumaru H., Maruyama A., Murakami H., Hosoyama A., Kaito C.,
Mizutani U., Takahashi N. K., Sawano T., Inoue R., Kaito C.,
Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C.,
Shiba T., Hattori M., Ogasawara N., Hayashi H. and Hiramatsu K.
Whole genome sequencing of methicillin-resistant Staphylococcus
aureus
Lancet 357 (9264), 1225-1240 (2001)
JOURNAL MEDLINE
21311952
PUBMED 11418146
REFERENCE 2 (bases 1 to 346900)
Ohta, T.
Direct Submission
Submitted (28-FEB-2001) Toshiko Ohta, University of Tsukuba College
of Medical Technology and Nursing, Department of Medical
Technology; 1-1-1 Ten-nodai, Tsukuba, Ibaraki 305-8577, Japan
(E-mail:tohta@akura.cc.tsukuba.ac.jp, Tel:81-298-53-3454,
Fax:81-298-53-3454)
COMMENT On May 29, 2001 this sequence version replaced gi:13875626.
FEATURES
Location/Qualifiers
1..346900
/organism="Staphylococcus aureus subsp. aureus Mu50"
/mol_type="genomic DNA"
/strain="Mu50"
/sub_species="aureus Mu50"
/db_xref="taxon:158878"
12..263
/gene="SAV1313"
12..263
/gene="SAV1313"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAB57475.1"
/db_xref="GI:14247084"
/translation="MLTLKLEDEQVLIYEIPEDDISNGKSVTFNKDKAEVIDFS
LSEINSEETFMLYRNKSFVSVRDFIEKQSFENYKIAYI"
987..1181
/gene="SAV1314"
987..1181
/gene="SAV1314"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAB57476.1"
/db_xref="GI:14247085"
/translation="MNLQSYIGKLVKLTLTNKILGKVIDPDDKVDNFDGYSNIEID
TGRITVDISENKITVLQK"
complement(1725..2750)
/gene="SAV1315"
complement(1725..2750)
/gene="SAV1315"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAB57477.1"
/db_xref="GI:14247086"
/translation="MISFENDYLEGAHERKVLRLVDNLVQASGYGDFQTAQAIKIK
KTDICPNATIRFLVGGTQTNQVINSMESEYGVISADTGHVAVHEGGAIEYSGHKV
IITPSKEGVSADSVETMETFKSPDKDMVPGMVIYISHPEYGTLYSKSELELC
KVCKQYQLPMDGRLGGLMSDQSDWTIKDIKCYCDVFIIGGTGIGALCGEAIYFT
KNNPKQFTTRIKHHGALLAKRLTGIQFLFELTDLNLYFNISRHALLEMANOKDGFRN
KGYRLYDFPTNQFFILSEKIALEQKFAVWEKYDDQHRVVRFAFATSWATTEENL
NKLELI"
3020..3217
/gene="SAV1316"
3020..3217
/gene="SAV1316"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAB57478.1"


```

/db_xref="GI:14247087"
/translation="MEQIKLITFTAETLLELLEKKNINAFLLSSEATNLKLVNITIKIEI
BTFPNNEEPNAILTSLNK"
gene
3238..4755
/gene="SAV1317"
CDS
3238..4755
/gene="SAV1317"
/codon_start=1
/transl_table=11
/product="cardiolipin synthetase homolog"
/protein_id="BAB57479.1"
/db_xref="GI:14247088"
/translation="MFFVPPFICEKTMRYTFNSDLGTLFTIILAIFINLVLAFLII
FLERNRRASSWAMLFVLFLGFLILYFFGRVTSARKLNKNGNVLTFDGLLK
QIESFDKNGYQKNDKQVOKHDLVSLMLMDQDGLTENNKVDFHFDGNDLVQVLDK
IKNAKEYHLEYVTFALDGLGRILHAEELKQGLEVKLYDDVGSKNVKNANPDHF
KSLGSEVAFPSAKPLPLNFRNRRNHRKLIIVDQQLGVGGFNGDSEYLGKLGKYG
RUTHRIQDAVALQRLFDLWNSQARHPQDEYDVKYPKNGPLGNSPQIAASGP
ASDWHQIEGYTKTMSAKKSYLOSPFIPDINSYNAIKIAKSGVDVHLMIPCKPD
HPLVWTFPSNADLLSSGVKIYTYENGPIHSMKCLIDDEIVSVGTAMDFRPFELN
EYNAFYDENLAKDLRVAYEHDITKSKQLTKESYANRPLSVKFKESLAKLVSPIL"
gene
4939..5436
/gene="SAV1318"
CDS
4939..5436
/gene="SAV1318"
/note="ATP-binding protein"
/codon_start=1
/transl_table=11
/product="ABC transporter homolog"
/protein_id="BAB57480.1"
/db_xref="GI:14247089"
/translation="MIQISNINKSNFKRCVLKNIISFDIEQKCIALIGKAGKSTLI
DILIGNVANSGEIPDKKLLQSENRSIMFOKTPPDQKVIEIINLYQSFYENPLPL
REIIEITKPSQLNQFYNKLSGGQRLDLFVLSLIGQPLILLDEPTSTWDIEIREY
FWSII"
gene
5461..5838
/gene="SAV1319"
CDS
5461..5838
/gene="SAV1319"
/note="ATP-binding protein"
/codon_start=1
/transl_table=11
/product="ABC transporter homolog"
/protein_id="BAB57481.1"
/db_xref="GI:14247090"
/translation="WLYTSHYIEEVERMSDKILLIENGCEIILNDSTSHRTWQSQIIT
LSDEYIRKLKDKDLVIQKNANGIKIITSNVNDTILYLOQLHINLDDIEIQKVSIV
DSYFNKKQKQSGNYDTKLEENRI"
gene
5807..6538
/gene="SAV1320"
CDS
5807..6538
/gene="SAV1320"
/note="similar to ABC transporter integral membrane
protein"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAB57482.1"
/db_xref="GI:14247091"
/translation="MILSYLKIEFKVIMRKKTLLLSILFPVIFILFTSILELPEDV
KPKFYKMYMSMTVYSLLSFLTFPLDIINEKQNEWRQRLMVPFTFTSYYSIKVVK
TMQFAIALIVFMVGHFYKGVAMSAVOWLESQIFLWLGASLLITFGILFSLINDIQK
TSALANIVLTIGLAVLGLGWFPINTPENMLQHVAVLPSYHLRKLGVDIASNNHILIS
FAILLVLGSLIAYVICISHPKRAE"
gene
6542..7633
/gene="SAV1321"
CDS
6542..7633
/gene="SAV1321"
/note="similar to two-component sensor histidine kinase"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAB57483.1"

```

```

/db_xref="GI:14247092"
/translation="MKFLKQTSIAEISSILYLIFPIAGIFFENEVYGPKWLYIISVIVF
SLSYLILVNNRLNTLMFYILLIHHVFLICVFPVSVHPMLSLFFFAFAPVPTFKN
NVKKTATNLFILTMILITILYLYNNYFVAMVYVYVVISLIMLNDPKMKNRKYOKE
IAEKRNHINTLIAQERHRICQDLHDLCHVPASLSLSELAYKLIDTDVKEKRAELL
AINKLSRESLNKVRERIDDKVLPSPFIEBISIRKVLKDDADIDFTFENKELAQVLSPTK
QSMVMTITREAINNVIKHANASKVHGKLTNNHKLKLLMIEDDGKIDSDCEVKSISQ
RVOLNGLTAVDSTNGTKIITEISTGGIA"
gene
7630..8232
/gene="SAV1322"
CDS
7630..8232
/gene="SAV1322"
/note="similar to two-component response regulator"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAB57484.1"
/db_xref="GI:14247093"
/translation="MTSLIIAEDQNLQRQAMVQILKHGDFEILADTDNGLDAMKLEI
EYNPNVILDIEMPCMTGLEVLARIKKHLNIRKVIIVTTPKPGYFEKAVVNDVAYV
LKRSIELVETINKVNGEKEYSATLMTSPFDKNPLTPKEQIVLREIGNGLSSKEI
SEKFLDTGTVRNVTSVIDKLADNRFDAMKANERKGI"
gene
8672..9205
/gene="nuc"
CDS
8672..9205
/gene="nuc"
/note="SAV1323"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAB57485.1"
/db_xref="GI:14247094"
/translation="MNISNENTVLIIVIAIIIGIFLQIFPKLPLIVTAVLSILLGI
FVGFTVIVIVSPYNNKRN"
gene
8672..9205
/gene="nuc"
CDS
8672..9205
/gene="nuc"
/note="SAV1324"
/codon_start=1
/transl_table=11
/product="thermonuclease"
/protein_id="BAB57486.1"
/db_xref="GI:14247095"

```

Query Match 100.0%; Score 2201; DB 1; Length 346900;
Best Local Similarity 100.0%; Pred. No. 5.4e-308;
Matches 2201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	TAATTTAAAAAGCAACTATTGTATAGAAAAATACAAAAATTTAAAAATATATATTACCTTATTA	60
DB	111604	TAATTTAAAAAGCAACTATTGTATAGAAAAATACAAAAATTTAAAAATATATATTACCTTATTA	111545
QY	61	GAATAAGTCGTATATAGAGGTGTACAAATGACGCAAAATTTTAATAGTAGAAGATGAACAA	120
DB	111544	GAATAAGTCGTATATAGAGGTGTACAAATGACGCAAAATTTTAATAGTAGAAGATGAACAA	111485
QY	121	AACCTAGCAAGATTTCTTGAATTGGAACCTCACACATGAAAAATTACAAATGTGGACACAGAG	180
DB	111484	AACCTAGCAAGATTTCTTGAATTGGAACCTCACACATGAAAAATTACAAATGTGGACACAGAG	111425
QY	181	TATGATGCAAGACGCGTTTATAGATAAAGCGCTTAGCCATTACTATGATTTAATCATATTA	240
DB	111424	TATGATGCAAGACGCGTTTATAGATAAAGCGCTTAGCCATTACTATGATTTAATCATATTA	111365
QY	241	GATTTAATGTTGCGCTCAATTAATGGCTTAGAAATTTGCGCAAAATTTAGACAAACAA	300
DB	111364	GATTTAATGTTGCGCTCAATTAATGGCTTAGAAATTTGCGCAAAATTTAGACAAACAA	111305
QY	301	TCTACACCTATCATTTATATATTACACGCAAAAGTGTAGTATGACAAAGTTGCTGGGCTT	360
DB	111304	TCTACACCTATCATTTATATATTACACGCAAAAGTGTAGTATGACAAAGTTGCTGGGCTT	111245
QY	361	GATTACGGTGCAGACGATTATATAGTTTAAGCGGTTTGTATTTAGAAAGACTTTTAGCAAGA	420

111244	Db	 GATTACGGTGCAGACGATATATAGTTAAGCCGTTTGATATTGAAGAACTTTTAGCAAGA	1111815
421	Qy	 ATTCGTGCAATTTTACGTGCGTCAGCCACAAAAAGGATATTATCGATGCTCAACGSTATTA	480
111184	Db	 ATTCGTGCANATTTTACGTGCGTCAGCCACAAAAAGGATATTATCGATGCTCAACGSTATTA	111125
481	Qy	 ATTGATAAGAACGCTTTTAAAGTGACGGTAAATGCGCGCAAGAAATTTGAAATTAACAAAAACA	540
111124	Db	 ATTGATAAGAACGCTTTTAAAGTGACGGTAAATGCGCGCAAGAAATTTGAAATTAACAAAAACA	111065
541	Qy	 GAGTATGATTTACTATATCTTCTAGCTGAAATATAAAACCATGTTATGCAACGGGAACAA	600
111064	Db	 GAGTATGATTTACTATATCTTCTAGCTGAAATATAAAACCATGTTATGCAACGGGAACAA	111005
601	Qy	 ATTTTAAATCATGTATGGGTTTATATAGTGAAGTAGAAACAAATGCTCGTAGATGTTTAT	660
111004	Db	 ATTTTAAATCATGTATGGGTTTATATAGTGAAGTAGAAACAAATGCTCGTAGATGTTTAT	110945
661	Qy	 ATAAGATATTTACGAAACAAAGTTAAACCATACGATCGTGACAAAATGATGAAACAGTT	720
110944	Db	 ATAAGATATTTACGAAACAAAGTTAAACCATACGATCGTGACAAAATGATGAAACAGTT	110885
721	Qy	 CGTGGCGTTGGGTATGTGATACGATGACAAAAACGTAATGCGCAATAACTGGAATATTG	780
110884	Db	 CGTGGCGTTGGGTATGTGATACGATGACAAAAACGTAATGCGCAATAACTGGAATATTG	110825
781	Qy	 TTACCACCATGATACGTTTGTGACGATATTTTCTGTTGTTTAAATATTTATTTTCT	840
110824	Db	 TTACCACCATGATACGTTTGTGACGATATTTTCTGTTGTTTAAATATTTATTTTCT	110765
841	Qy	 TGAAGATACACTGTCATTAATAGTGAAGTTGATGATGCAAGCAAGCTCAACGCGATATTA	900
110764	Db	 TGAAGATACACTGTCATTAATAGTGAAGTTGATGATGCAAGCAAGCTCAACGCGATATTA	110705
901	Qy	 ATAATTTATTTCTTAAGCCTGTTTAAAGATATATCTGCATTAGACTTTGAATGCATCTT	960
110704	Db	 ATAATTTATTTCTTAAGCCTGTTTAAAGATATATCTGCATTAGACTTTGAATGCATCTT	110645
961	Qy	 TAGGTAAATTTCAAGAGATATTTATTTATGATGAGCATATTAATTAATTTTGAAGACT	1020
110644	Db	 TAGGTAAATTTCAAGAGATATTTATTTATGATGAGCATATTAATTAATTTTGAAGACT	110585
1021	Qy	 CGAATGATAACACAGTGAGAGTTGAACACAGGTTATGAACACCGTTATTTTGACCGGTAA	1080
110584	Db	 CGAATGATAACACAGTGAGAGTTGAACACAGGTTATGAACACCGTTATTTTGACCGGTAA	110525
1081	Qy	 TAAAAAAACGCTATAAAGCAATTTGAATATTTAAATTTTAAAGAACCAATTAACACGCAAG	1140
110524	Db	 TAAAAAAACGCTATAAAGCAATTTGAATATTTAAATTTTAAAGAACCAATTAACACGCAAG	110465
1141	Qy	 ATTTTCAAGGGTATAGCTTGTTATTTCAATCTACTAGAAAATTTATGATACATCGTAAAT	1200
110464	Db	 ATTTTCAAGGGTATAGCTTGTTATTTCAATCTACTAGAAAATTTATGATACATCGTAAAT	110405
1201	Qy	 CATTTGATATCATTTGCGCTGGCATTTGGAGTGATTGCAACCAATTAATGCGCAACAATCA	1260
110404	Db	 CATTTGATATCATTTGCGCTGGCATTTGGAGTGATTGCAACCAATTAATGCGCAACAATCA	110345
1261	Qy	 GTTATGATATTTTCAACACAAATTTACTAAACCGCTTTGTCAGTTTATCAAAATAAATGATTG	1320
110344	Db	 GTTATGATATTTTCAACACAAATTTACTAAACCGCTTTGTCAGTTTATCAAAATAAATGATTG	110285
1321	Qy	 AGATTCGACGAGATGGTTTTCCAAATTAATTTGCAATTAATTAACAAATTTATGAGAAATAG	1380
110284	Db	 AGATTCGACGAGATGGTTTTCCAAATTAATTTGCAATTAATTAACAAATTTATGAGAAATAG	110225
1381	Qy	 ATAATTTAGCAAAATACGTTTAAATGAGATGATGAGCAAAATTTGAAGAAATCATTTATCAAC	1440
110224	Db	 ATAATTTAGCAAAATACGTTTAAATGAGATGATGAGCAAAATTTGAAGAAATCATTTATCAAC	110165
1441	Qy	 AAAGACAATTTGTTGAAGATCGGTCACATGAATTTACGAACACCAATTAACAAATTTATCAAG	1500

Db	110164	AAAGA	CAATTTT	GTG	AAGAT	CGCT	CA	CA	TGA	TTAC	GAAC	AC	CA	TTAC	GAAT	TAT	TC	CAAG	110105	
Qy	1501	GTCA	TTTAAAT	TTG	ATTC	CAG	CGAT	GGG	GA	AAAA	AG	CC	CAG	CA	GTAT	TAG	AA	GAAT	CGT	1560
Db	110104	GTCA	TTTAAAT	TTG	ATTC	CAG	CGAT	GGG	GA	AAAA	AG	CC	CAG	CA	GTAT	TAG	AA	GAAT	CGT	110045
Qy	1561	TAAAT	TATTT	CTAT	TG	AG	AAAT	TGA	AT	TC	GTAT	CA	TAAA	AT	TAG	TG	CA	GAAT	TACT	1620
Db	110044	TAAAT	TATTT	CTAT	TG	AG	AAAT	TGA	AT	TC	GTAT	CA	TAAA	AT	TAG	TG	CA	GAAT	TACT	109985
Qy	1621	TGAC	TAAAG	GAG	AT	GT	AAAT	TG	AC	ATTT	CT	TC	TG	AA	GCG	GAG	AC	CGT	GC	1680
Db	109984	TGAC	TAAAG	GAG	AT	GT	AAAT	TG	AC	ATTT	CT	TC	TG	AA	GCG	GAG	AC	CGT	GC	109925
Qy	1681	AAAT	TCGT	CCG	GA	TAC	AC	TC	ATT	TAA	AA	CA	AT	TC	GC	AT	TAT	CA	AT	1740
Db	109924	AAAT	TCGT	CCG	GA	TAC	AC	TC	ATT	TAA	AA	CA	AT	TC	GC	AT	TAT	CA	AT	109865
Qy	1741	AT	CT	GAC	AT	CT	AAAA	AT	CT	GA	AAAT	T	AAAA	T	GAA	AC	CT	CAT	CAA	1800
Db	109864	AT	CT	GAC	AT	CT	AAAA	AT	CT	GA	AAAT	T	AAAA	T	GAA	AC	CT	CAT	CAA	109805
Qy	1801	TAA	CTTT	TAT	TG	AT	TG	CA	AT	T	CA	AA	T	AT	GA	TG	GA	AA	AT	1860
Db	109804	TAA	CTTT	TAT	TG	AT	TG	CA	AT	T	CA	AA	T	AT	GA	TG	GA	AA	AT	109745
Qy	1861	CAAG	TTAAAA	AAAT	TAAG	CA	AAAA	TAAT	TG	AAAT	T	TAC	AG	AT	CAT	CG	GAAT	T	CG	1920
Db	109744	CAAG	TTAAAA	AAAT	TAAG	CA	AAAA	TAAT	TG	AAAT	T	TAC	AG	AT	CAT	CG	GAAT	T	CG	109685
Qy	1921	AGGA	GAT	CA	AG	AT	T	TC	AT	T	T	T	T	T	T	T	T	T	T	1980
Db	109684	AGGA	GAT	CA	AG	AT	T	TC	AT	T	T	T	T	T	T	T	T	T	T	109625
Qy	1981	GTCA	AGG	CGG	TAA	TG	GA	CT	CG	GA	T	T	AT	TG	CT	CA	AAAA	AT	CA	2040
Db	109624	GTCA	AGG	CGG	TAA	TG	GA	CT	CG	GA	T	T	AT	TG	CT	CA	AAAA	AT	CA	109565
Qy	2041	GAT	CG	AT	T	AAAA	T	T	AAAA	TG	CA	AA	T	T	AA	AG	CA	AC	CGT	2100
Db	109564	GAT	CG	AT	T	AAAA	T	T	AAAA	TG	CA	AA	T	T	AA	AG	CA	AC	CGT	109505
Qy	2101	CAT	GT	CT	GAG	AC	GT	CA	AT	T	CA	AA	G	T	CA	TAG	G	AT	CA	2160
Db	109504	CAT	GT	CT	GAG	AC	GT	CA	AT	T	CA	AA	G	T	CA	TAG	G	AT	CA	109445
Qy	2161	CTA	AT	G	T	A	T	A	G	A	A	C	A	C	T	A	T	A	A	2201
Db	109444	CTA	AT	G	T	A	T	A	G	A	A	C	A	C	T	A	T	A	A	109404

RESULT 4	
AF165314	
LOCUS	3083 bp DNA linear BCT 16-JUL-2000
DEFINITION	Staphylococcus aureus putative response regulator ArlR (arlR) and putative protein histidine kinase ArlS (arlS) genes, complete cds.
ACCESSION	AF165314
VERSION	AF165314.1 GI:9230551
KEYWORDS	.
SOURCE	Staphylococcus aureus
ORGANISM	Staphylococcus aureus
REFERENCE	Bacteria; Firmicutes; Bacillales; Staphylococcus. 1 (bases 1 to 3083) Fournier,B. and Hooper,D.C. A new two-component regulatory system involved in adhesion, autolysis, and extracellular proteolytic activity of Staphylococcus aureus
JOURNAL	J. Bacteriol. 182 (14), 3955-3964 (2000)
MEDLINE	20327575
PubMed	10869073
REFERENCE	2 (bases 1 to 3083) Fournier,B. and Hooper,D.C. Direct Submission
AUTHORS	
TITLE	

JOURNAL	Submitted (05-JUL-1999) Unite de Biochimie Microbienne, Institut Pasteur, 25 rue du Docteur Roux, Paris 75015, France
FEATURES	Location/Qualifiers
source	1..3083
	/organism="Staphylococcus aureus"
	/mol_type="genomic DNA"
	/strain="ISP794"
gene	/db_xref="taxon:1280"
	336..995
CDS	/gene="arlR"
	336..995
	/gene="arlR"
	/codon_start=1
	/transl_table=11
	/product="putative response regulator ArlR"
	/protein_id="AAF85896.1"
	/db_xref="GI:9230552"
	/translation="MTQLIVDEQNARFLFLELTHENVNVDTEYDQDGLDKALSH YDILLDMLPSINGLEICRKIRQQSPPIIIITAKSDTYDKVAGLOYGADDDYIVKP FDIELLARIRAIRLRQPKQDIIDVNGITIDKNAPKVTWNGAEIELTETEDLLYLLA ENKNHMOREQILNHWGYNSEVETNVVDVYIRYLRLNKLKPYDRDKMIETVRGVGYI R"
- gene	992..2347
	/gene="arlS"
CDS	992..2347
	/gene="arlS"
	/codon_start=1
	/transl_table=11
	/product="putative protein histidine kinase ArlS"
	/protein_id="AAF85897.1"
	/db_xref="GI:9230553"
	/translation="MTKRLRNWIIIVTMTFVIFLFLIIIFFLKDLHNSBLDD AERSSDINNLFHSPXVDISALDNLASLNGFOEIIIVDEHNKLFETISNDONTVRVEP GVEHYRFRVIRKRYKGYEYLIKEPIITQDPKGYSLLIHSLNVDNIVKSLIYIYALA PGVIATIIATISYVFQITKPLVSLNKMIEIRDFQONKLOLANTVREIDNLANT FNEWMSQIEESFNQRFQVEDASHLRPLQIIQGHNLIIQRWKKDPAYLEESLNIS IEMNRKILVEELLELTGVDNDISSEAGTQVHIINDEIRSHRSIHLUKLHPDYQFDIDL TSKNLEIKMKHQFQLFLIFIDNAIKYDVRNKKIKVTRLNKKQKIIETIDHGIGIP EEDQDFIEDRFRVDRKSRSRSGGNGLGLSIAQKIIQLNGSGSIKIKSEINKGTTFKII P"
ORIGIN	
Query Match	99.8%; Score 2196.2; DB 1; Length 3083;
Best Local Similarity	99.9%; Pred. No. 9.3e-307;
Matches 2198; Conservative	0; Mismatches 3; Indels 0; Gaps 0;
Qy	1 TAATTTAAAGCAACTATTGTAGAAAAATACAAAAATTTAAATATATACCTATTATA 60
Db	249 TAATTTAAAGCAACTATTGTAGAAAAATACAAAAATTTAAATATATACCTATTATA 308
Qy	61 GAAAAAGTCGTAAATAGAGTGTACAAATGACGCAAAATTTAATAGTAGAAGATGAACAA 120
Db	309 GAAAAAGTCGTAAATAGAGTGTACAAATGACGCAAAATTTAATAGTAGAAGATGAACAA 368
Qy	121 AACTTAGCAAGATTTCTTGAATTTGGAATCTCACATGAAAAATTAACAATGTGGACACAG 180
Db	369 AACTTAGCAAGATTTCTTGAATTTGGAATCTCACATGAAAAATTAACAATGTGGACACAG 428
Qy	181 TATGATGGACAGACGGTTAGATAAAGCGCTTAGCCATTACTATGATTTAATCATATTA 240
Db	429 TATGATGGACAGACGGTTAGATAAAGCGCTTAGCCATTACTATGATTTAATCATATTA 488
Qy	241 GATTTAAATGTCGCGTCAATTAATGGCTTAGAAAAATTTGCGCAAAATTAGACAAACAA 300
Db	489 GATTTAAATGTCGCGTCAATTAATGGCTTAGAAAAATTTGCGCAAAATTAGACAAACAA 548
Qy	301 TCTACACCTTATCATTATAATTACAGCGAAAAGTGATACGTATGACAAAGTTGCTGGGCTT 360
Db	549 TCTACACCTTATCATTATAATTACAGCGAAAAGTGATACGTATGACAAAGTTGCTGGGCTT 608
Qy	361 GATTACGGTGCAGACGATTATATAGTTAAGCCGTTTGATTTGAAGAAGCTTTTAGCAAGA 420
Db	609 GATTACGGTGCAGACGATTATATAGTTAAGCCGTTTGATTTGAAGAAGCTTTTAGCAAGA 668

Qy	421 ATTGCTGCAATTTTACGTCGTGAGCCACAAAGGATATTTATCGATGTCAACGGTATTACA 480
Db	669 ATTGCTGCAATTTTACGTCGTGAGCCACAAAGGATATTTATCGATGTCAAAGGATTTACA 728
Qy	481 ATTGATAGAAGCGCTTTTAAAGTACGCTTAATGCGCGAGAAATTTGAATTAACAAAAACA 540
Db	729 ATTGATAGAAGCGCTTTTAAAGTACGCTTAATGCGCGAGAAATTTGAATTAACAAAAACA 788
Qy	541 GAGTATGATTTACTATATCTTCTAGCTGAAAAATAAAACCATGTTATGCAACGGGAACAA 600
Db	789 GAGTATGATTTACTATATCTTCTAGCTGAAAAATAAAACCATGTTATGCAACGGGAACAA 848
Qy	601 ATTTTAAATCATGATGGGTTATAATAGTGAAGTAGAAAACAAAATGTCGTAGATGTTTAT 660
Db	849 ATTTTAAATCATGATGGGTTATAATAGTGAAGTAGAAAACAAAATGTCGTAGATGTTTAT 908
Qy	661 ATAAGATATTTACGAACAAGTTAAAACCATACATCGTGACAAAAATGATTTGAAACAGTT 720
Db	909 ATAAGATATTTACGAACAAGTTAAAACCATACATCGTGACAAAAATGATTTGAAACAGTT 968
Qy	721 CGTGCCTGGGTATGTGATACGATGACAAACCGTAAATTCGCAATTAACCTGGATTATTG 780
Db	969 CGTGCCTGGGTATGTGATACGATGACAAACCGTAAATTCGCAATTAACCTGGATTATTG 1028
Qy	781 TTACCACGATGATTACGTTTGTCTACGATATTTTGTGTTTGTGTTTAAATTTATTTTCT 840
Db	1029 TTACCACGATGATTACGTTTGTCTACGATATTTTGTGTTTGTGTTTAAATTTATTTTCT 1088
Qy	841 TGAAGATACACTGCAATTAATAGTACGTTGATGAGCAAGCAAGCTCAAGCGATATTA 900
Db	1089 TGAAGATACACTGCAATTAATAGTACGTTGATGAGCAAGCAAGCTCAAGCGATATTA 1148
Qy	901 ATAAATTTATTTTCAAGCTGTAAAGATATATCTGCAATAGCATTTGAATGATCTT 960
Db	1149 ATAAATTTATTTTCAAGCTGTAAAGATATATCTGCAATAGCATTTGAATGATCTT 1208
Qy	961 TAGTAAATTTTCAAGAGATAATTTATGATGAGCATAATAATAAATTTATTTGAGACAT 1020
Db	1209 TAGTAAATTTTCAAGAGATAATTTATGATGAGCATAATAATAAATTTATTTGAGACAT 1268
Qy	1021 CGAATGATACACAGTGAAGTTGAAACGAGTTATGAACACCGTTATTTTGACCGCGTAA 1080
Db	1269 CGAATGATACACAGTGAAGTTGAAACGAGTTATGAACACCGTTATTTTGACCGCGTAA 1328
Qy	1081 TAAAAAAGCGCTATAAAGCGATTTGAATTTTAAAGAACCAATTTCAACGCAAG 1140
Db	1329 TAAAAAAGCGCTATAAAGCGATTTGAATTTTAAAGAACCAATTTCAACGCAAG 1388
Qy	1141 ATTTCAAAGGGTATAGCTTTGTTAAATTCATTCACCTAGAAAAATTTATGATTAACATCGTAAAT 1200
Db	1389 ATTTCAAAGGGTATAGCTTTGTTAAATTCATTCACCTAGAAAAATTTATGATTAACATCGTAAAT 1448
Qy	1201 CATTTGATATCATTTGCGCTGGCATTTGGAGTGTGCAACAAATTAATACTGCGCAACATCA 1260
Db	1449 CATTTGATATCATTTGCGCTGGCATTTGGAGTGTGCAACAAATTAATACTGCGCAACATCA 1508
Qy	1261 GTTATGATTTTCAACACAAATTTACTAAACCGCTTGTCAAGTTTCAATTAATGATGTTG 1320
Db	1509 GTTATGATTTTCAACACAAATTTACTAAACCGCTTGTCAAGTTTCAATTAATGATGTTG 1568
Qy	1321 AGATTGACGAGATGGTTTTTCAAAATAAATTTGCAATTTAAATACAAAATTTATGAAGAAATAG 1380
Db	1569 AGATTGACGAGATGGTTTTTCAAAATAAATTTGCAATTTAAATACAAAATTTATGAAGAAATAG 1628
Qy	1381 ATAATTTAGCAATAACGTTTAAATGAGATGATGAGCGCAAAATTTGAAGAAATCATTTTATCAAC 1440
Db	1629 ATAATTTAGCAATAACGTTTAAATGAGATGATGAGCGCAAAATTTGAAGAAATCATTTTATCAAC 1688
Qy	1441 AAGCAATTTGTTGAGATCGGTACATGAATTTACGAACACCATTTACAAATTTATTCAAG 1500
Db	1689 AAGCAATTTGTTGAGATCGGTACATGAATTTACGAACACCATTTACAAATTTATTCAAG 1748


```

/gene="MW1100"
/note="ORFID:MW1100
hypothetical protein, similar to RNA-binding Sun protein"
/codon_start=1
/transl_table=11
/protein_id="BAB94965.1"
/db_xref="GI:21204268"
/translation="MIENRSLAFDTIIDIINEGAYSNLRINEVLSSENLNMDKALF
TEIVGTGKRFKFTLDYFKPFVKTKIKAWRQLWMSYQVYLDKVNPHAIINEAVE
IAKERGGVHNGVNIILRTWMSDLDPENIADPKKMALEYSPKWIIDHMAITHYG
LEKTEITLQSELETSTTVRANLFRASLDLDIIEKLQDEGYDVEKDHDLPCYLHGGQP
IHKRSFGKGFVSQKSMFVAHIMNVRDHRDVLDAACAFGGKACHIAEVLMPGGQV
DASDIHDKILINFIKIKRLTNIKAFOHDATPKYDKTIDKILVDAPCSGLGWRHK
PEIKYTKQKHISELVELQLELENKNNVIGGIEIYSTCTIEQLENNVIYTFLKN
NKNFEFPQHPITGELVKTLQIMPQDENSDFITIKRKDN"
4642..5736
/gene="MW1101"
4642..5736
/gene="MW1101"
/note="ORFID:MW1101"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAB94966.1"
/db_xref="GI:21204269"
/translation="MITAEKKKKKFLPNFKQSIYSRFDQMQLVVEQQQKFRPK
QIFEWLYQKRVDSIDEMTNLSKDLQQLKDNFTVTTLTVVKQSKOGTIFKLYELOQ
GYTETLVRMDYNGNSCVTTQVCGRICTFCASLTGLGKLENLEAGEIVSVLTVOKA
LDATTEERSQVIMIGIGEPFENYDEMDFLRIVNDNSLNIGARHITVSTSGIIPRIY
LPADEDIQINFAVSLHAARDEVRSLRPNIRAYINVEKLEIAIYQYQKTNRRVIFEY
LFGVQDQLEHARELAHLIIGLNCHVNLIIPVNHVERNYVKTAKNDIPEKFEKELKRLG
INATIRREQSGSIDAACQLRAKERQVETR"
5743..6486
/gene="MW1102"
5743..6486
/gene="MW1102"
/note="ORFID:MW1102"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAB94967.1"
/db_xref="GI:21204270"
/translation="MLEAQFFDTQGRDKNEDAGGIIFYNRINQQLLVLCDMGSHKA
GEVASKFTDELKSRFEAEALIEEHOAENLRNNIKDINFOLYHVAQENASYKMGTT
CVCALFVSQVIANVNGDSRAVINSROIQIOTSDHSFVNHLVLGTQITPREAPTHQ
RNILTKMGTDKRVSPDLFIKELNEFYDVLNLSGLTDLYVKDNEIKRLLVKEGFIEDH
GDQLMQLALDNHNSKDNVTFLAAIEGDKV"
6483..8477
/gene="MW1103"
6483..8477
/gene="MW1103"
/note="ORFID:MW1103
hypothetical protein, similar to protein kinase"
/codon_start=1
/transl_table=11
/protein_id="BAB94968.1"
/db_xref="GI:21204271"
/translation="MIGKILNERYKIVDKLGGGNGSTVYLAEDTILNIKVAIKFIP
PREKETUKRFERVHNSQSHQNIKSMIDVEDDCYLLVMETIEGPTLSEYIESH
GPLSDVAINFTQVQIQLGDKHAHDMRIVHRDIKPNILIDSNKTLKIFDFGIKALSE
TSLTQTNHVLTVQVYFSEPOAKGEATDECTDIYSIGIVLYEMLVGEPPFNGETAVSIA
IKHIQDSVPNVNTDVRKIDPOSLSNVILRATKDKANRYKTIQEMKDDLSSVLHENRA
NEDVYELDKMTIAVPLAKEDIKHISEKSNQPKRETTOVPIVNGPAAHQFOKQPEG
TYPEPKKSKTRKVLVLSLFLSMIALVSPVANMFGNKYEETPDVIGKSVKEAQ
IFNKNLKLKGISRSOKIPENELIKTTPNGERVERGSDVVVISGSPKRVKMPNV
IGLPKEALQKLSLGLKDVTKIEKYNNAQPKGYTANQSVTANTEIAIHDSNLIKYES
LGIQVYVEDFEHKSFSKAKALEEKGFKVESDDIDEGDVISQSPKGSVDEG
STISFVSGKKSDDSDVKVTTTESVDVPTVTKNDKSKQKVYIKDKDNDGSTEKGSFD
ITSQORIDIPILRIEKGKTASYIVKVDGKTVAEKSVDV"
8705..9580
/gene="MW1104"
8705..9580
/gene="MW1104"

```

```

/note="ORFID:MW1104"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAB94969.1"
/db_xref="GI:21204272"
/translation="MKTGRIVSISGVYQVDVNGERFNTKPRGLFRKKKFSPPVGVDIV
EFQVQNEGYIHQVFERKNEKRPVPSNIDTLVIMSAPENFSQLLDRFLVIAHS
YQLNARVLVTKTKPIEKQFIEINELIKYENIYETEYEFIGNDDRRKKIVRAWPAGLI
VLSQSGVGKSTFLNHYRPELNLTNDISKSLNRGKHTTRHVELPERQNGYIADTPGF
SALDPDHDIKDEIKDYFLELNRGETCKFRNCNHIKEPNCNVKQLEIGNIAQFRYDH
YLFQFNEISNRKVRV"
9581..10225
/gene="cfx8"
9581..10225
/gene="cfx8"
/note="ORFID:MW1105
ribulose-5-phosphate 3-epimerase homolog"
/codon_start=1
/transl_table=11
/protein_id="BAB94970.1"
/db_xref="GI:21204273"
/translation="MTKLYPSLLSVDFLDLQHELKRLLEEAGVDGVHEDVMDGQVFPNI
SIGLPILDVAVRKGTTLFDVHLMIEENPEKTIASFAERGADMISIHVSTPHIRAIOM
IKHLDKAGVYINPFTPISQIEPILDIVDYLVTMVPNGFGQSFIDQCVEKIAQLNA
IKMERQNFIDIEVDGGVNTDTAKVCVENGATMLVTGSPFPKQEDYKKVTQALKG"
10232..10873
/gene="MW1106"
10232..10873
/gene="MW1106"
/note="ORFID:MW1106"
/codon_start=1
Query Match 99.2%; Score 2183.4; DB 1; Length 295350;
Best Local Similarity 99.5%; Pred. No. 1.9e-305; Indels 0; Gaps 0;
Matches 2190; Conservative 0; Mismatches 11;
Qy 1 TAAATTTAAAAAGCAACTATTCTATAGAAAAATACAAAAATTTAAAAATATTACCTTATTA 60
Db 231341 TAAATTTAAAAAGCAACTATTCTATAGAAAAATACAAAAATTTAAAAATATTACCTTATTA 231282
Qy 61 GAAAAAGTCGTAATATGAGGTGTCAAAATGACGCAAAATTTTAAATAGTAGAAGATGAAACAA 120
Db 231281 GAAAAAGTCGTAATATGAGGTGTCAAAATGACGCAAAATTTTAAATAGTAGAAGATGAAACAA 231222
Qy 121 AACTTAGCAAGATTTCTTGAAATTCGAACTCACATCAATGAAATTTCAATGTGACACACAG 180
Db 231221 AACTTAGCAAGATTTCTTGAAATTCGAACTCACATGAAATTTCAATGTGACACACAG 231162
Qy 181 TATGATGACAAAGACGGTTTAGATAAAGCGCTTAGCCATTACTATGATTTAATCATATTA 240
Db 231161 TATGATGACAAAGACGGTTTAGATAAAGCGCTTAGCCATTACTATGATTTAATCATATTA 231102
Qy 241 GATTAAATGTTGCGGTCAATTAATGCGTTAGAAAATTTGTGCGCAAAATTTAGACAAACAA 300
Db 231101 GATTAAATGTTGCGGTCAATTAATGCGTTAGAAAATTTGTGCGCAAAATTTAGACAAACAA 231042
Qy 301 TCTACACCTATCATTAATTTACACGCAAAAGTCATAGTATGACAAAGTTGCTGGGCTT 360
Db 231041 TCTACACCTATCATTAATTTACACGCAAAAGTCATAGTATGACAAAGTTGCTGGGCTT 230982
Qy 361 GATTACGGTCGACACGATTATATAGTTAAGCCGTTTGTATTTGAAGAACTTTTATAGCAAGA 420
Db 230981 GATTACGGTCGACACGATTATATAGTTAAGCCATTTGTATTTGAGAAGACTTTTAGCAAGA 230922
Qy 421 ATTGCGTCAATTTTACGTCGTGACCCACAAAGGATATTATCGATGTCACCGTATTACA 480
Db 230921 ATACGTGCGATTTTACGTCGTGACCCACAAAGGATATTATCGATGTCACCGTATTACA 230862
Qy 481 ATTGATAGAAGCGCTTTTAAAGTCACGTTAATGCGCGCAGAAATTTGAATTAACAAAACA 540
Db 230861 ATTGATAGAAGTCGCTTTTAAAGTCACGTTAATGCGCGCAGAAATTTGAATTAACAAAACA 230802

```

QY 541 GAGTATGATTATCTATATCTTTCTAGCTGAAATATAAAACCAATGTTATGCAACGGGAACAA 600
Db 230801 GAGTATGATTACTATATCTTTCTGCTGAAATATAAAACCAATGTTATGCAACGGGAACAA 230742
QY 601 ATTTTAAATCATGATGCGGTTTAAATAGTGAAGTAGAACAATAATGTCGTAGATGTTTAT 660
Db 230741 ATTTTAAATCATGATGCGGTTTAAATAGTGAAGTAGAACAATAATGTCGTAGATGTTTAT 230682
QY 661 ATAAGATATTATTCGAACAACAGTTAAACCATACCATGCTGACAAAATGATTGAACACAGTT 720
Db 230681 ATAAGATATTATTCGAACAACAAATTAACCATACCATGCTGACAAAATGATTGAACACAGTT 230622
QY 721 CGTGGCGTGGGTATGTGATACGATGACAAACGTAATTCGCGCAATAACTGGAATTATG 780
Db 230621 CGTGGCGTGGGTATGTGATACGATGACAAACGTAATTCGCGCAATAACTGGAATTATG 230562
QY 781 TTACACGATGATTACGTTTGTGACGATATTTTTGTTTGTGTTTAAATTAATTTTCT 840
Db 230561 TTACACGATGATTACGTTTGTGACGATATTTTTGTTTGTGTTTAAATTAATTTTCT 230502
QY 841 TGAAGATATACACTGCATAATAGTCAGCTTGTATGATGACGAACGAAGCTCAAGCGATATTA 900
Db 230501 TGAAGATATACACTGCATAATAGTCAGCTTGTATGATGACGAACGAAGCTCAAGCGATATTA 230442
QY 901 ATAAATTTATTTCAATCTAAGCTGTTAAAGATATATCTGCAATTAGACTTGAATGCATCTT 960
Db 230441 ATAAATTTATTTCAATCTAAGCTGTTAAAGATATATCTGCAATTAGACTTGAATGCATCTT 230382
QY 961 TAGTAAATTTCAAGAGATTAATTTATGATGAGCATATAATAATAATTTTGAAGCAT 1020
Db 230381 TAGTAAATTTCAAGAGATTAATTTATGATGAGCATATAATAATAATTTTGAAGCAT 230322
QY 1021 CGAATGATAACACAGTGCAGAGTTCAACACAGTTATGAACACCGTTATTTTGAACCGGTAA 1080
Db 230321 CGAATGATAACACAGTGCAGAGTTGAACACAGTTATGAACACCGTTATTTTGAACCGGTAA 230262
QY 1081 TAAAAAACGCTATAAAGGCATTGAATATTTAAATTTAAAGAACCAATTACAACGCAAG 1140
Db 230261 TAAAAAACGCTATAAAGGCATTGAATATTTAAATTTAAAGAACCAATTACAACGCAAG 230202
QY 1141 ATTTCAAAGGTTAGCTTGTAAATTCATTCATCAGTAAATTTATGATAAATCGTAAAT 1200
Db 230201 ATTTCAAAGGTTAGCTTGTAAATTCATTCATCAGTAAATTTATGATAAATCGTAAAT 230142
QY 1201 CATTGATATCATTCGCTGCTGCTTTCGAGTGATGCAACAAATTAATCACTGCCACATCA 1260
Db 230141 CATTGATATCATTCGCTGCTGCTTTCGAGTGATGCAACAAATTAATCACTGCCACATCA 230082
QY 1261 GTTATGTTATTTCAACACAAATTAATAACCGCTTGTGAGTTATCAAAATAAATGATTG 1320
Db 230081 GTTATGTTATTTCAACACAAATTAATAACCGCTTGTGAGTTATCAAAATAAATGATTG 230022
QY 1321 AGATTCCAGAGATGTTTTCAAAATAAATGCAATTAATTAATACAAATTTAGAAATAG 1380
Db 230021 AGATTCCAGAGATGTTTTCAAAATAAATGCAATTAATTAATTAATAGAAATAG 229962
QY 1381 ATAATTTAGCAATACGTTTAAATGAGATGATGAGCAAAATTTGAAGATCATTTAATCAAC 1440
Db 229961 ATAATTTAGCAATACGTTTAAATGAGATGATGAGCAAAATTTGAAGATCATTTAATCAAC 229902
QY 1441 AAGACAATTTGTTGAAGATGCGTCACATGAATTAACGAACCAATTAACAAATTTATCAAG 1500
Db 229901 AAGACAATTTGTTGAAGATGCGTCACATGAATTAACGAACCAATTTATCAAG 229842
QY 1501 GTCATTTAAATTTGATTACCGATGGGGAATAAAGCCAGACGATTTAGAAATCGT 1560
Db 229841 GTCATTTAAATTTGATTACCGATGGGGAATAAAGCCAGACGATTTAGAAATCGT 229782
QY 1561 TAAATTTCTTATTGAAGAATGAATCGTATCATAAATTTAGTCGAAGATTTACTTGAAT 1620
Db 229781 TAAATTTCTTATTGAAGAATGAATCGTATCATAAATTTAGTCGAAGATTTACTTGAAT 229722
QY 1621 TGACTAAAGGAGATGTTAAATGACATTTCTTCTGAAGCGCAGACCGTGCAATTAATGATG 1680

Db 229721 TGACTAAAGGAGATGTTAAATGACATTTCTTCTGAAGCACAGACGTCATATTAATGATG 229662
QY 1681 AAATTCGCTCGCAATACACTCATTAATAAACAATTCGATTCGATTAATCAATTTGATACGG 1740
Db 229661 AAATTCGCTCGCAATACACTCATTAATAAACAATTCGATTCGATTAATCAATTTGATACGG 229602
QY 1741 ATCTGACATCTAAAAATCTAGAAATTAATAATGAACCTCATCAATTCGAACAATTTATTT 1800
Db 229601 ATCTGACATCTAAAAATCTAGAAATTAATAATGAACCTCATCAATTCGAACAATTTATTT 229542
QY 1801 TAATCTTTATTTGATTAATGCAATCAAAATATGATGTAAGATAAGAAAATTTAAAGTTAAAGA 1860
Db 229541 TAATCTTTATTTGATTAATGCAATCAAAATATGATGTAAGATAAGAAAATTTAAAGTTAAAGA 229482
QY 1861 CAAGTTTAAAAATTAAGCAAAAATAATTTGAATTAACAGATCATGGAATTTGTTCCAG 1920
Db 229481 CAAGTTTAAAAATTAAGCAAAAATAATTTGAATTAACAGATCATGGAATTTGTTCCAG 229422
QY 1921 AGAAGATCAAGATTTTCAATTTTGTGCTTTTATCGAGTGAATAATCTCGTTCAAGAA 1980
Db 229421 AGAAGATCAAGATTTTCAATTTTGTGCTTTTATCGAGTGAATAATCTCGTTCAAGAA 229362
QY 1981 GTCAAGCGGTAAATGGAATTCGATTCATTTCTCAAAAAATCAATTAATAACGGAG 2040
Db 229361 GTCAAGCGGTAAATGGAATTCGATTCATTTCTCAAAAAATCAATTAATAACGGAG 229302
QY 2041 GATCGATTAATAATTAAGTGAATAATAAGGAACAAGTTTAAATCATATTTTAAAT 2100
Db 229301 GATCGATTAATAATTAAGTGAATAATAAGGAACAAGTTTAAATCATATTTTAAAT 229242
QY 2101 CATGCTGAGACGTCATCAATCAAGTCATAGATCAATTTTAAAGTACACATTAAGTGTGA 2160
Db 229241 CATGCTGAGACGTCATCAATCAAGTCATAGATCAATTTTAAAGTACACATTAAGTGTGA 229182
QY 2161 CTAATGTATAAGAACCACTATAAAACAAATAAACAAGTGT 2201
Db 229181 CTAATGTATAAGAACCACTATAAAACAAATAAACAAGTGT 229141

RESULT 6
AP003133/c

LOCUS

DEFINITION AP003133 303750 bp DNA linear. BCT 11-JAN-2003
Staphylococcus aureus subsp. aureus N315 genomic DNA, complete
genome, section 5/10.

ACCESSION

AP003133 BA000018

VERSION

AP003133.2 GI:14349175

KEYWORDS

Staphylococcus aureus subsp. aureus N315

SOURCE

Bacteria; Firmicutes; Bacillales; Staphylococcus.

ORGANISM

1

REFERENCE

AUTHORS

Kuroda, M., Ohta, T., Uchiyama, I., Baba, T., Yuzawa, H., Kobayashi, I.,

Cui, L., Oguchi, A., Aoki, Y., Nagai, Y., Lian, J., Ito, T., Kanamori, M.,

Mizutsumaru, H., Maruyama, A., Murakami, H., Hosoyama, A.,

Matsutani, U. Y., Takahashi, N. K., Sawano, T., Inoue, R., Kaito, C.,

Sekimizu, K., Hirakawa, H., Kuhara, S., Goto, S., Yabuzaki, J.,

Kanehisa, M., Yamashita, A., Oshima, K., Furuya, K., Yoshino, C.,

Shiba, T., Hattori, M., Ogasawara, N., Hayashi, H. and Hiramatsu, K.

Whole genome sequencing of methicillin-resistant Staphylococcus

aureus

Lancet 357 (9264), 1225-1240 (2001)

21311952

11418146

2 (bases 1 to 303750)

Director-General, Biotechnology Center, Aoki, K., Oguchi, A.,

Hosoyama, A., Nagai, Y., Kuroda, M., Hiramatsu, K. and Kikuchi, H.

Direct Submission

Submitted (30-JAN-2001) Director-General, Biotechnology Center,

National Institute of Technology and Evaluation, Biotechnology

Center; 2Chome 49-10 Nishihara, Shibuya-ku, Tokyo 151-0066, Japan

(E-mail: biocenter.go.jp, URL: http://www.bio.nite.go.jp/.

Tel: 81-3-3481-1933, Fax: 81-3-3481-8424)

COMMENT On Jun 12, 2001 this sequence version replaced gi:13701012.

```
FEATURES
    Location/Qualifiers
        1..303750
            /organism="Staphylococcus aureus subsp. aureus N315"
            /mol_type="genomic DNA"
            /strain="N315"
            /sub_species="aureus N315"
            /db_xref="taxon:158879"
        202..1155
            /gene="SA1056"
        202..1155
            /gene="SA1056"
            /note="ORFID:SA1056"
            /codon_start=1
            /transl_table=11
            /product="hypothetical protein"
            /protein_id="BAB42308.1"
            /db_xref="GI:13701013"
            /translation="MKRTIGCLLLIMLVAGCSFGNGHLSKKSEESKQETVKES
            EEEKDPLEKYEIEIKMKGIKDAPSLDKLPLMTKFTNSKGIQGWKDYKELMGKV
            ELADYRFTKDSKSSIKDVDAFFKGKGIKRKVIETHDDVKQDVYVDPDGKIGNS
            NTPVYAIMTKYKDGKLVASVEPGSVIHKODAIKYDDYSLKSLQLTKLHDPKP
            VPYSAQIKSGVPLTSVPMTHGSKDTKDEVLPALEYTFSPRYEDKSNPDPKVLN
            LVHMDPLNASSDFGNAHFVVLJSKYIKYESNYETASDLSLK"
            complement(1318..1521)
            /gene="SA1057"
            /note="SA1057"
            /note="ORFID:SA1057"
            /codon_start=1
            /transl_table=11
            /product="conserved hypothetical protein"
            /protein_id="BAB42309.1"
            /db_xref="GI:13701014"
            /translation="MQVKNYIYLLALFLSGFIHKFYAGKNMQGILHLIFCWGIPH
            ILAIIASAVITVPKPADEQGNVTL"
            1820..2308
            /gene="SA1058"
            /note="SA1058"
            /note="ORFID:SA1058"
            /product="hypothetical protein, similar to polypeptide deformylase"
            /codon_start=1
            /transl_table=11
            /protein_id="BAB42310.1"
            /db_xref="GI:13701015"
            /translation="MAIKLVPASHPIITKAQAVKTFDDSLKRLQLDLEMTVAQEA
            AGLCAPQNSLQVAIIDMEGLQLVNPKIISQSNETITDLEGSITLDPVYGEVTR
            SKMIVESYDVGNGKVELTAHEDVARMILHIIDQNGIPPTERADRIITRKEVAYFI
            ND"
            2301..3236
            /gene="SA1059"
            2301..3236
            /gene="SA1059"
            /note="ORFID:SA1059"
            /codon_start=1
            /transl_table=11
            /product="methionyl-tRNA formyltransferase"
            /protein_id="BAB42311.1"
            /db_xref="GI:13701016"
            /translation="MTKIIIPMGTPDFTTIVLEMLIAEHADVIAVVTQDPVPGRKRVMT
            PPVKVAMKYDLVPYQPKLSGSELEQLQLVDVLTAAFGQLLPESLLALPKLG
            AINVASILPKYRGAPTHOAIIDGQEQTGITMYMVKLDAGNIISQAIKIRENDN
            VGTWHDKLSVLGADLLKETPLSIIIEGTNESVPQDQTQATPASNIRREDERISWPKPCR
            QVFNQIRGLSPFWVAYITMDINLKIYDAELVETNKINEPGTILBTIKAILVATNDN
            EAVATKQWLQAGKMLAANYLSGAQNTLVGKKLI"
            3155..4540
            /gene="SA1060"
            3155..4540
            /gene="SA1060"
            /note="ORFID:SA1060"
            /product="hypothetical protein, similar to RNA-binding Sun protein"
            /codon_start=1
```

```

            /transl_table=11
            /protein_id="BAB42312.1"
            /db_xref="GI:13701017"
            /translation="MSWEKENVSCQLFKKCAKHTSRRETYMIENVRSIAPDTIODILN
            EGAYSINRINEVLSENELNAMDKAFTIETVYGTVKRKYTLDFYLPKPFVKTIKAWVRQ
            LLWMSIYQVYLDKPNHAIINEAVEIAKRGVTHGNGVNVNGIILKTMRSOLDPNEI
            ADPKRMAIEYSPMKPIIDHWATHYGLEETILOSFLFETTTVRANLTRASLDSDII
            EKLQDEGVDEKHDLPYCLHIGQPPIHSRSFKDGFYSIQKSMFVAHITMNVDRHD
            HVLDCAPGKACHIAEVLMPGEGVDASDIHDHKLIDLINEFNIKKLRTNIIKAFQHDA
            TKPYDKTYDKILVDAPCSGLGMRHKPEIKYTQSKOHIESLVELQLEILENVKNVKI
            GGEIISTCTIEQLEENVIYFLKNNKPFEPFQHPITGELVYKTLQIMPQDQNSDG
            FFIITKIRKKN"
            4543..5637
            /gene="SA1061"
            4543..5637
            /gene="SA1061"
            /note="ORFID:SA1061"
            /codon_start=1
            /transl_table=11
            /product="conserved hypothetical protein"
            /protein_id="BAB42313.1"
            /db_xref="GI:13701018"
            /translation="MITAEKKKKNKPLPNFDKQSIYSLRPFDMQNLVQCOQKFRK
            QIFEWLYOKRVDSIDEMTNLSKQLQLKDNFTVTLTVVKQSKDGTIKPLFELQD
            GYTETVLMRHDYGNVCVTQVGCRICTFCASTLGLKKNLEAGEIVSQLVQKA
            LDATERVSVQIVIMGIEPPEFENYDEMDFLRIVNDNSLNIGARHITVSTSRVTPIRY
            DPADEDIQINFVSLHAAKDEVRSLMPINRAYNVKLEIAIOYQETNRVRVFEYG
            LFGVNDQLEHARELAHLIKGLCHGNLIPVNHVPERNVYVTKANDIPKFEKELKRLG
            INATIRREQSDIDAACQLRAKERQVETR"
            5644..6387
            /gene="SA1062"
            5644..6387
            /gene="SA1062"
            /note="ORFID:SA1062"
            /codon_start=1
            /transl_table=11
            /product="conserved hypothetical protein"
            /protein_id="BAB42314.1"
            /db_xref="GI:13701019"
            /translation="WLSAQFFDTGQHRDKNEDAGGI FYNOTNOQLLVLCQWGGHKA
            GEVASFVTDLKSFRPAENLIEHQAEWLNRNNIKDINFOLYHYAQENAEYKMGMTT
            CVCALVFEKSVIVIANVGSRAIVINSQIEQITSDHSFVNLVLTGQITPBEAFTHPQ
            RNIITKVMTGDKRVSPLFIKRLNFYDLLNSDGLTDYVKNDEIKRLLVKEGTIEDH
            GDQLMQLALDNHSDNVTFILAAIEGDKV"
            6384..8378
            /gene="SA1063"
            6384..8378
            /gene="SA1063"
            /note="ORFID:SA1063"
            /codon_start=1
            /transl_table=11
            /product="protein kinase"
            /protein_id="BAB42315.1"
            /db_xref="GI:13701020"
            /translation="MIGKIINERYKIVDKLGGGGMSTVYLAEDTILNIKVAIKAFIP
            PREKEETLKRFEVHNSOLSQHNIVSMIDVEDDCYLVMEVIEGPTLSYEYTESH
            GPLSDVTAINTNOILDGIKHAHDMRIVHRDIKPNILIDSNKTLKIFDGIKALSE
            TSLQTNHVLGTQVYFSPQKAGTDECTDIYSIGIVLYEMLVGEPPFNGETASVIA
            IKHIQDSVPNVTTDVRKDIPOSNLIRATEKDKANKYTIQEMKDDLSVSLHENRA
            NEDYVDEPKTKTIAVPLKEDLAKHISEKSNQPKRETQVPIVNGPAHQOQFQEQ
            TVYEPKPKKTRKIVLLSLIFSLMIALSVFAMAMFGNKBETPDVIGSKVKEAEQ
            IFGNKLQKLSRSYSDKYPENIIKTPTNGRVERGSDVDVVISKGPBKVMNPV
            IGLPKEALOKLKSGLKDVITIEKVYNOAPKGVIANQSVTANTEIAIHDSNLIKLYES
            LGIQVTVFEPEHKSFSKAKALEKGFVESKEYSDDIDEGDVISSGPKGKSVDEG
            STISFVSGKKSDDVKTITTESVDVPTGKNDKSKQKVYIITDKONDGSTEKGSFD
            ITSDQRIDIPLRIEKGTASYIVKVDGKTVAEKEVSVDII"
            8606..9481
            /gene="SA1064"
            8606..9481
            /gene="SA1064"
            /note="ORFID:SA1064"
            /codon_start=1
            /transl_table=11
```


/product="conserved hypothetical protein"	
/protein_id="BAB42316.1"	
/db_xref="GI:13701021"	
/translation="MKTRIVKISGVYVDVNGERNTKPRGLFRKKKSPVVDIV	
EFVQNIETGHIQVFRKQKELKPPVSNIDTLVMSAVEPNFSTQLDRFLVIAHS	
YQLARVLVTKKQTPLEKQFENELKLIYENIGYETEFIQNDGDRKKIYVAVPAGLI	
VLSQSGVGKSTFLNHRPELNDISLNRKHTLRHVELFRQNGYIADTPGF	
SALDFDIDKDEIKDYFLELNRNYGETCKFRNCNHIKEPNCNVKHQLEIGNIAQPRYDH	
YLOLFNEISNRKRY"	
gene	9482..10126
/gene="cfxb"	
CDS	9482..10126
/gene="cfxb"	
/note="ORFID:SA1065	
ribulose-5-phosphate 3-epimerase homolog"	
/codon_start=1	
/transl_table=11	
/protein_id="BAB42317.1"	
/db_xref="GI:13701022"	
/translation="MTKLVPSSLVSDFLDLQHELKRLLEEAGVDGVHVDVMDGQFVPI	
SIGLPILDVVRKGTLLPIDVHLMLENPEKYIASFAEHGADMISHVESTPHIRAQM	
IKHLDKAGVIVNGTPISQIEPILDIVLVNTPNPGGGGSDFDQCVEKLAGLNA	
IKMERQNFLEVDGGVNTDTAKVCVENGATMLVTGSPFKQSDYKRVTOALKG"	
Query Match 98.1%; Score 2159; DB 1; Length 303750;	
Best Local Similarity 99.3%; Pred. No. 6.2e-302;	
Matches 2185; Conservative 0; Mismatches 0; Indels 16; Gaps 1;	
Qy	1 TAAITTTAAAAGCAACTATTGTTATAGAAAAATACAAAATTTAAAAATATATTACCTTATTA 60
Db	228274 TAAITTTAAAAGCAACTATTGTTATAGAAAAATACAAAATTTAAAAATATATTACCTTATTA 228215
Qy	61 GAAAAGTCGTAATAAGAGGTGTAACAATGACGCAATTTTAAATAGTAGAGATGACAA 120
Db	228214 GAAAAGTCGTAATAAGAGGTGTAACAATGACGCAATTTTAAATAGTAGAGATGACAA 228155
Qy	121 AACTTAGCAAGATTTCTTGAATGGAACTCACACATGAAAATTTAAATGTGGACACAGAG 180
Db	228154 AACTTAGCAAGATTTCTTGAATGGAACTCACACATGAAAATTTAAATGTGGACACAGAG 228095
Qy	181 TATGATGGACAGACGGTTAGATAAAGCGTTAGCCATTACTATGATTTAAATCATATTA 240
Db	228094 TATGATGGACAGACGGTTAGATAAAGCGTTAGCCATTACTATGATTTAAATCATATTA 228035
Qy	241 GATTTAATGTTCGCGTCAATTAATGGCTTAGAATTTGTGCGCAAAATTTAGACACACAA 300
Db	228034 GATTTAATGTTCGCGTCAATTAATGGCTTAGAATTTGTGCGCAAAATTTAGACACACAA 227975
Qy	301 TCTACACCTATCAATTAATTAAGCGAAAGTGATACGTATGACAAAGTTGCTGGGCTT 360
Db	227974 TCTACACCTATCAATTAATTAAGCGAAAGTGATACGTATGACAAAGTTGCTGGGCTT 227915
Qy	361 GATTAACGGTCAGACGATTATATAGTTAAGCGTTTGATATGAAAGCTTTTAGCAAGA 420
Db	227914 GATTAACGGTCAGACGATTATATAGTTAAGCGTTTGATATGAAAGCTTTTAGCAAGA 227855
Qy	421 ATTGCTGCAATTTTACGTCGTCAGCCACAAAGGATATTATCGATGTCACGCGTATTACA 480
Db	227854 ATTGCTGCAATTTTACGTCGTCAGCCACAAAGGATATTATCGATGTCACGCGTATTACA 227803
Qy	481 ATTGATAAGAACGCTTTTAAAGTGACGGTAAATGGCGCAGAAATTTGAATTTAAACAAACA 540
Db	227802 -----GAACGCTTTTAAAGTGACGGTAAATGGCGCAGAAATTTGAATTTAAACAAACA 227751
Qy	541 GAGTATGATTACTATATCTTCTAGCTGAAATTAATAAACAATGTTATGCAACGGGAACAA 600
Db	227750 GAGTATGATTACTATATCTTCTAGCTGAAATTAATAAACAATGTTATGCAACGGGAACAA 227691
Qy	601 ATTTTAAATCATGTATCGGGTTTAAATAGTGAAGTGAACAAACAATGTCGTAGATGTTTAT 660
Db	227690 ATTTTAAATCATGTATCGGGTTTAAATAGTGAAGTGAACAAACAATGTCGTAGATGTTTAT 227631
Qy	661 ATAAGATATTACGAAACAAGTTAAACCCATACGATCGTGACAAAATGATTGAAACAGTT 720
Db	227630 ATAAGATATTACGAAACAAGTTAAACCCATACGATCGTGACAAAATGATTGAAACAGTT 227571
Qy	721 CGTGGCGCTTGGGTATGTGATACGATGACAAAACGTAATTTGCGCAATTAACCTGGAATATTG 780
Db	227570 CGTGGCGCTTGGGTATGTGATACGATGACAAAACGTAATTTGCGCAATTAACCTGGAATATTG 227511
Qy	781 TTACCCAGATGATTACGTTTGTCTACGATATTTTGTGTTGTTGTTTAAATTTATTTTCT 840
Db	227510 TTACCCAGATGATTACGTTTGTCTACGATATTTTGTGTTTGTGTTTAAATTTATTTTCT 227451
Qy	841 TGAAGATACACTGCTAATAATAGTGAGCTTGATGATGACAGAACGAACTCAAGCGATATTA 900
Db	227450 TGAAGATACACTGCTAATAATAGTGAGCTTGATGATGACAGAACGAACTCAAGCGATATTA 227391
Qy	901 ATAAATTTATTTCACTTCTAAGCCTGTTAAAGATATATCTGATAGCTTGAATGCACTCT 960
Db	227390 ATAAATTTATTTCACTTCTAAGCCTGTTAAAGATATATCTGATAGCTTGAATGCACTCT 227331
Qy	961 TAGTAAATTTCAAGAGATAATTTATTTATGATGAGCATATAATAATTTTGGAGCAT 1020
Db	227330 TAGTAAATTTCAAGAGATAATTTATTTATGATGAGCATATAATAATTTTGGAGCAT 227271
Qy	1021 CGAATGATACACAGTCAGAGTTGAAACCGGTTATGAACACCGTTATTTGACCGCGTAA 1080
Db	227270 CGAATGATACACAGTCAGAGTTGAAACCGGTTATGAACACCGTTATTTGACCGCGTAA 227211
Qy	1081 TAAAAAACCCTATAAAGGCAITTTGAATATTTAAATTTAAAGAACCAATTTACACGCAAG 1140
Db	227210 TAAAAAACCCTATAAAGGCAITTTGAATATTTAAATTTAAAGAACCAATTTACACGCAAG 227151
Qy	1141 ATTTCAAGGATAGCTTGAATTTCACTTCACTAGAAATTTATGATGATGATGATGATGAT 1200
Db	227150 ATTTCAAGGATAGCTTGAATTTCACTTCACTAGAAATTTATGATGATGATGATGATGAT 227091
Qy	1201 CATTGATATCATTTGGCGTGGCAATTTGGAGTGATGCAACAAATTAATTAACCTGCCACATCA 1260
Db	227090 CATTGATATCATTTGGCGTGGCAATTTGGAGTGATGCAACAAATTAATTAACCTGCCACATCA 227031
Qy	1261 GTTATGATTTTCAACACAAATTTAACCCTGTCAGTTTCAAAATTAATTAATTAATTAAT 1320
Db	227030 GTTATGATTTTCAACACAAATTTAACCCTGTCAGTTTCAAAATTAATTAATTAATTAAT 226971
Qy	1321 AGATTGACGAGATGTTTCAAAATTAATTTGCAATTTAAATTAATTAATTAATTAATTAAT 1380
Db	226970 AGATTGACGAGATGTTTCAAAATTAATTTGCAATTTAAATTAATTAATTAATTAATTAAT 226911
Qy	1381 ATAAATTTAGCAATACGTTTAAATGAGATGATGACCAAAATTTGAAGATCAATTTAATCAAC 1440
Db	226910 ATAAATTTAGCAATACGTTTAAATGAGATGATGACCAAAATTTGAAGATCAATTTAATCAAC 226851
Qy	1441 AAAGCAATTTGTTGAAGATGCGTCACATGAATTTAGAACCAATTTACAAATTTATTTCAAG 1500
Db	226850 AAAGCAATTTGTTGAAGATGCGTCACATGAATTTAGAACCAATTTACAAATTTATTTCAAG 226791
Qy	1501 GTCAATTTAAATTTGATTCAGCGATGGGAAAAAGACCCAGCAGTATTAGAGAGATCGT 1560
Db	226790 GTCAATTTAAATTTGATTCAGCGATGGGAAAAAGACCCAGCAGTATTAGAGAGATCGT 226731
Qy	1561 TAAATTTCTATTGAAAGAAATGAATCGTATCATATAAATTTAGTCGAAGAAATTTACTTGAAT 1620
Db	226730 TAAATTTCTATTGAAAGAAATGAATCGTATCATATAAATTTAGTCGAAGAAATTTACTTGAAT 226671
Qy	1621 TGACTAAAGAGATGTAATGATCATTTCTTCTGAAGCGCAGACCGTGCATATTAATGATG 1680
Db	226670 TGACTAAAGAGATGTAATGATCATTTCTTCTGAAGCGCAGACCGTGCATATTAATGATG 226611
Qy	1681 AAATTCGCTCGGAAATACACTCATTAATAAACAATTTGATCTCTGATTTCAATTTGATACGG 1740
Db	226610 AAATTCGCTCGGAAATACACTCATTAATAAACAATTTGATCTCTGATTTCAATTTGATACGG 226551
Qy	1741 ATCTGATCTAAAAATCTAGAAATTTAAATGAAACCTCATCAATTTGCAACAAATTTATTT 1800

Db	226550	ATCTGCATCTAAAAATCTAGAAATTAATAATGAACCTCATCAATTCGAACAAATATTTT	226491	Db	244	TAGAAATTTGTCGCAAAATTTAGACAACAACAAATCTACACCTATCATTTATTAATTACAGCGA	303																				
Qy	1801	TAATCTTTATTGATTAATGCAATCAAAATATGATGTGAAGAAATTAAGGTTAAGA	1860	Qy	329	AAAAGTGATACGTATGACAAAAGTTGCTGGCTTGATTACGGTGCAGACGATTAATATAGTTA	388																				
Db	226490	TAATCTTTATTGATTAATGCAATCAAAATATGATGTGAAGAAATTAAGGTTAAGA	226431	Db	304	AAAAGTGATACGTATGACAAAAGTTGCTGGCTTGATTACGGTGCAGACGATTAATATAGTTA	363																				
Qy	1861	CAAGGTTAAAAAATTAAGCAAAAAATTAATTTGAAATTTACAGATCATGGAATTTGGTATTCAG	1920	Qy	389	AGCCGTTTGATTTAGAAGAACTTTTAGCAAGAAATTCGTGCAATTTTACGTCGTGAGCCAC	448																				
Db	226430	CAAGGTTAAAAAATTAAGCAAAAAATTAATTTGAAATTTACAGATCATGGAATTTGGTATTCAG	226371	Db	364	AGCCGTTTGATTTAGAAGAACTTTTAGCAAGAAATTCGTGCAATTTTACGTCGTGAGCCAC	423																				
Qy	1921	AGGAGATCAAGATTTTCAATTTTGTATCGCTTTTATCGAGTGGATAAAATCTCGTTCAAGAA	1980	Qy	449	AAAAGGATTAATTCGATGTCAACGGTATTACAATTTGATAAGAACGCTTTTAAAGTGACGG	508																				
Db	226370	AGGAGATCAAGATTTTCAATTTTGTATCGCTTTTATCGAGTGGATAAAATCTCGTTCAAGAA	226311	Db	424	AAAAGGATTAATTCGATGTCAACGGTATTACAATTTGATAAGAACGCTTTTAAAGTGACGG	483																				
Qy	1981	GTCAAGCGGTAATGGACTCGGATTTATCTATTGCTCAAAAAATCATTCAAATTAACCGGAG	2040	Qy	509	TAAATGGCGCAGAAATTTGAATTTAAACAAAACAGAGTATGATTTACTATATCTTCTAGCTG	568																				
Db	226310	GTCAAGCGGTAATGGACTCGGATTTATCTATTGCTCAAAAAATCATTCAAATTAACCGGAG	226251	Db	484	TAAATGGCGCAGAAATTTGAATTTAAACAAAACAGAGTATGATTTACTATATCTTCTAGCTG	543																				
Qy	2041	GATCGATTAAAAATTAAGAGTGAATTAATAAAGGAACAACGTTTAAATCATATTTTAAT	2100	Qy	569	AAAAATAAAACCATTTGATGCAACGGGAACAAAATTTTAAATCATGTATGGGGTTATAATA	628																				
Db	226250	GATCGATTAAAAATTAAGAGTGAATTAATAAAGGAACAACGTTTAAATCATATTTTAAT	226191	Db	544	AAAAATAAAACCATTTGATGCAACGGGAACAAAATTTTAAATCATGTATGGGGTTATAATA	603																				
Qy	2101	CATGCTGAGAGCTCAATCAAAAGTCATAGGATCAATTTTAAAGTACACATTTAGCTGTGA	2160	Qy	629	GTGAAGTAGAAAAAATGTCGTAGATGTTTATATATAAGATATTTTACGGAACAAGTTTAAAC	688																				
Db	226190	CATGCTGAGAGCTCAATCAAAAGTCATAGGATCAATTTTAAAGTACACATTTAGCTGTGA	226131	Db	604	GTGAAGTAGAAAAAATGTCGTAGATGTTTATATATATAGRTATTTTACGGAACAAGTTTAAAC	663																				
Qy	2161	CTAATGTATAGAACAACCTATATAAACAAATAAACAGTGGTT	2201	Qy	689	CATACGATCGTGACAAAAATGATTCAAAACAGTTTCGTGGCTTGGGTATGTGATACGATGAC	748																				
Db	226130	CTAATGTATAGAACAACCTATATAAACAAATAAACAGTGGTT	226090	Db	664	CATACGATCGTGACAAAATGATTCAAAACAGTTTCGTGGCTTGGGTATGTGATACGATGAC	723																				
RESULT 7																											
LOCUS	AR354607	1440 bp	DNA	linear	PAT 17-AUG-2003	Qy	749	AAAAACGTAATTTGGCGCAATAA	808																		
DEFINITION	Sequence 725 from patent US 6593114.							Db	724	AAAAACGTAATTTGGCGCAATAA	783																
ACCESSION	AR354607							Qy	809	ATTTTGTGTTTGTGTTTAAATTTATTTTCTTCAAAGATACACATGCAATATAGTGAGCT	868																
VERSION	AR354607.1	GI:33760691						Db	784	ATTTTGTGTTTGTGTTTAAATTTATTTTCTTCAAAGATACACATGCAATATAGTGAGCT	843																
KEYWORDS	Unknown.							Qy	869	TGATGATGACAGAACGAAGCTCAAGCGATATTAATAATTTATTTTCATTTCTAAGCCCTGTTAA	928																
SOURCE	Unknown.							Db	844	TGATGATGACAGAACGAAGCTCAAGCGATATTAATAATTTATTTTCATTTCTAAGCCCTGTTAA	903																
ORGANISM	Unknown.							Qy	929	AGATATATCTGCATTAGACTTTGAATGCAATCTTTAGGTAATTTTCAAGAGATAAATTTATTA	988																
REFERENCE	1 (bases 1 to 1440)							Db	904	AGATATATCTGCATTAGACTTTGAATGCAATCTTTAGGTAATTTTCAAGAGATAAATTTATTA	963																
AUTHORS	Kunsch, C.A., Choi, G.H., Barash, S., Dillon, P.J., Fannon, M.R. and Rosen, C.A.							Qy	989	TGATGAGCATTAATAATAATTTATTTTGAGACATCGAATGATAACACAGTGAGATTGAACC	1048																
TITLE	Staphylococcus aureus polynucleotides and sequences							Db	964	TGATGAGCATTAATAATAATTTATTTTGAGACATCGAATGATAACACAGTGAGATTGAACC	1023																
JOURNAL	Patent: US 6593114-A 725 15-JUL-2003;							Qy	1049	AGGTTATGAACACCGTTATTTTGACCGGTAAATAAATAACGCTATATAAGGCATTGAATA	1108																
FEATURES	Location/Qualifiers							Db	1024	AGGTTATGAACACCGTTATTTTGACCGGTAAATAAATAACGCTATATAAGGCATTGAATA	1083																
source	i. .1440							Qy	1109	TTTAAATTTATAAGAACCAATTTACAAACGCAAGATTTTCAAAAGGGTATAGCTTGTAAATTC	1168																
ORIGIN	/organism="unknown"							Db	1084	TTTAAATTTATAAGAACCAATTTACAAACGCAAGATTTTCAAGGGTATAGCTTGTAAATTC	1143																
								Qy	1169	TTCACTAGAAAAATTAATGATAAATCATCGTTAAATTCATTTGCGCTGGCAATTTGG	1228																
								Db	1144	TTCACTAGAAAAATTAATGATAAATCATCGTAAATTCATTTGTATATCATTTGCGCTGGCAATTTGG	1203																
								Qy	1229	AGTGATTCGCAACAATTAATGATGCGCAACAATCAGTTATGTAATTTTCAACACAAATTTACTAA	1288																
								Db	1204	AGTGATTCGCAACAATTTAATGATGCGCAACAATCAGTTATGTAATTTTCAACACAAATTTACTAA	1263																
								Qy	1289	ACCGCTTGTAGTTTATCAAAATAAATGATTTGAGATTCGACGAGATGGTTTCAACAAATAC	1348																
								Db	1264	ACCGCTTGTAGTTTATCAAAATAAATGATTTGAGATTCGACGAGATGGTTTCAAAATAA	1323																
								Qy	1349	ATTGCAATTAATAACAAATTTATGAAGAAATAGATAAATTTAGCAAAATCGTTTAAATGAGAT	1408																


```
/mol_type="genomic DNA"
/strain="ATCC 12228"
/db_xref="ATCC:12228"
/db_xref="taxon:176280"
113..346
/locus_tag="SE0907"
113..346
/locus_tag="SE0907"
/transl_table=11
/product="acyl carrier protein"
/protein_id="AAO04504.1"
/db_xref="GI:27315370"
/translacion="MENFDKVDIIIVRLGVADKVTDASFKDDLGADSLDIAELVM
ELEDFGTETPDEEAEKINTVGDAVKYINSLEK"
461..1198
/locus_tag="SE0908"
461..1198
/locus_tag="SE0908"
/transl_table=11
/product="RNase III"
/protein_id="AAO04505.1"
/db_xref="GI:27315371"
/translacion="MANQKKEMVHNFQOKFTDKMSLGLHFKNIDLYQQAFSHSPFI
NDFNNRLEHNERLEFLGDAVLELTVSRYLPHLPEGNLTKMRATIVCEPSLVIF
ANKILNELILGKEEKTGTRPSLISDAFEAFVGLYLDQGLDSVMTAEKVIIP
VVEDLGVVDFTQFOEYVHSONKGVTVQLIKEGPAHRLFTSEVILENKAAVE
GKGTKESEQAQAQYAKMKKKSL"
1325..4894
/locus_tag="SE0909"
1325..4894
/locus_tag="SE0909"
/transl_table=11
/product="chromosome segregation SMC protein"
/protein_id="AAO04506.1"
/db_xref="GI:27315372"
/translacion="WVYLKSIDAIGFSAFADHTNVQDKGVTAIVPGNSGSKSNITDA
IKWLGEOSAKSLRCARMEDII FSGAEHRKKAQNAEVLKLDNHSQKLIQDSEELVVT
RLRYSGSSEYLLNDRARLKDII DLFLDSGLGKEAFSII SQGRVDEILNAKPNRRQ
IIEESAGVSEYLLNDRARLKDII DLFLDSGLGKEAFSII SQGRVDEILNAKPNRRQ
LSKMEQSDIYVSDI DHYEDNQRDLERNHLKSQAQKEGQQAQQLQLQYKQ
RQNDYDEKLNYELVATENYQKSGVLNLEERKKQSETNAREBELDNLESQID
STKNEAKNELAELKQKOLNEVOLESLLYISDQHDEKLEETKNSYTYLMS
QSDVNDIRFLEHTINENAKSRDLSRLVEAFNOLKDIQQNITQTOKEYQSSKMS
KVEQNIQLEQLTDSKLLSEYENKLYQARYNEKLSRIDS LATQSEDIYFPNGV
KHLKAKDELRGIGHGVAEVINVPSEMTQAIETALGASLQHVIVDNEKGRQAIQYL
KORGGRATFLPLNVIQPRVAAADIKDVARGSQGFINTASDAINVSAYQNIENLLG
NTIIVENLKHANLARIYRTRIVTLEGDVVPNGSGMTGGARKTSILSOKDELST
MNQLEDYQROTAEPERQFKEQTOAEOLSEQYFSASQYNNLKEOVHHELELDRLK
TOEALHKNHEEFEEKNDGYQSDSKETLKEQNHLIEIQQLQLESDEIERYTQLS
KEGKASTQTQQLKQSDLAUVXERIKSQQVYERLDQLSDSERQKINEKIKL
FNSDMMGKADPELKEQIQOENVRQNLQOLSEIQQRDLNKEIKINESQLOKCH
QDILSIEHYQDIIKAKQSLDVIINHAI DNLNDTYQLTVARMEYDSDETDNLRRK
VKLTMSIDELGPLNLNAIEQPEELNERYTFLNEQRTDLRAKETLEQI IHENDKEVE
GRFTTHAVQDHFTTVFKOLFGGQAEHLRLTEDDYSAGVDIIIVPPGKKLOHLSLL
SGGERALNALLFALLKVRSAFPVILDEVEALDEANVIRYAOYLNELSTETQPIVI
THRGTMFSDRLGYVTQWESGVSKLVSNLNTIIDEVWKEQA"
4891..6117
/locus_tag="SE0910"
4891..6117
/locus_tag="SE0910"
/transl_table=11
/product="signal recognition particle"
/protein_id="AAO04507.1"
/db_xref="GI:27315373"
/translacion="MSFFKRLKDFSSKSEDDIQKDLDTVDNSNANSQSDGWDNDS
EQVKPKPKKLLSEADFEDGLISIEDFEETEAQKIGAKFAGKSKSQNQQLNLL
IARYKRVDEDFEALDEMLITADVGFTNWKLTDLBELTAQRNLQOETEDREVEIK
IVEIYHQEDDHSNAINEDGRNLVIMVGVGKTTTIGLAYRYQQEGKVMMAAG
DTFRAGAIQNLVWGERVGVVEVVSQNEGSDFAAVVYDAINAANKNGVDILICDTAGRL
```

```
gene
CDS

113..6451
/locus_tag="SE0911"
113..6451
/locus_tag="SE0911"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAO04508.1"
/db_xref="GI:27315374"
/translacion="MGQNDLLVKTURMNYLFDYQSOILLTNKKNLYLEFPFYLODYSLSSEI
ADTFEVRQAVYDNIIRRTGDLVEDYSEKRLYQRFKRELYNLKMSLNQSPPELLKQY
ITQLEELE"
6482..7849
/locus_tag="SE0912"
6482..7849
/locus_tag="SE0912"
/codon_start=1
/transl_table=11
/product="fff protein"
/protein_id="AAO04509.1"
/db_xref="GI:27315375"
/translacion="MAFEGLSDRLOQTQMKRGGKVTADIKTMREVRLLALLEADV
NFKVKGPFVKTSSERALGSDVMQSLTPQQVVIKI VQBELTSLMGCENTSIKVNKPPT
VWVVLQAGAKTTTAGLALLMRKKNKFLVAGDI RPAADIQLTQVGGQIDIPV
YSEGDVSPQOIVENALKHAKKEHLDFVIDTAGRLHIDEALMELQEVKEISKPDEI
MLVDMTQGDANVNAQSFQDLQVSGVTITKLOGDTRGGAAALSIRSTQPKFVGM
SEKLDGUELFHPERMASSILGMGDVLSLIEAKQDDVDQEKADLEKKNRESFTLDDF
LEQLDQVKNLGLDDIMKRI PGMNMGKGLDKLNSDKQIDHIKAI IQSMTPKERSNPE
TLNVSRRKRIAKGSGRSLSQEVNRLMKOFNDMKQWKQKTGCGKGKGAKNQWQNMKG
MNLPP"
8281..8556
/locus_tag="SE0913"
8281..8556
/locus_tag="SE0913"
/codon_start=1
/transl_table=11
/product="30S ribosomal protein S16"
/protein_id="AAO04510.1"
/db_xref="GI:27315376"
/translacion="MAVKIRLTRLGSKRNPFYRIVADARSPRDGRRIIEQLGTYNPAH
VNAPEVKIDEALAKLWLDGAKPTDTVHNILSREGILKKFDEQKNAK"
8683..9186
/locus_tag="SE0914"
8683..9186
/locus_tag="SE0914"
/codon_start=1
/transl_table=11
/product="rimM protein"
/protein_id="AAO04511.1"
/db_xref="GI:27315377"
/translacion="MEVEVGGQIVNTHGKGVKVSNSDFTETFPQGEQLLVKHNNT
IENVTVASRYIHKGHMLRFEGINNIEHLKGDYIYQERDHDQIELGEHEHYSDI
IGCTVFKDDTPIGRVINIPETGANDVWVVKEXYLIPYIEDVVKOIDIETKTIKIT
PMEGLD"
9186..9923
/locus_tag="SE0915"
9186..9923
/locus_tag="SE0915"
/codon_start=1
/transl_table=11
/product="tRNA (guanine-N1)-methyltransferase"
/protein_id="AAO04512.1"
/db_xref="GI:27315378"
/translacion="WKIDYLTLPFEMFEGVLNHSILKRAQDKGIINVTINFRDYSIN
KHNQVDYDFPGGQGVMLKPEPEVNAMEIDINNEHTRVILMCPQGRPTQIAOELSE
AKHTVLCGHVGEYDERIKRHLVTDIELSMGDYVLTGGELPAMTWTDATVRLIPGVIGN
QASHQDQSFSDGLLEFPQYTRPREYKMSVPEVLLSGNHAHIDQWRHEQKILRTYVEK
PDLLLEPLTDEKREILLETYYKKLUND"
10029..10379
/locus_tag="SE0916"
10029..10379
```

CDS	10029..10379 /locus_tag="SE0916" /codon_start=1 /transl_table=11 /product="50S ribosomal protein L19" /protein_id="AA004513.1" /db_xref="GI:27315379" /translation="MSNHKLI EAVTKSQRLTDLPTFRSGDTLRVHVRI VEGSRERIQV RSRVKRRGGISSETFTVRKISSGVGVERTFPLHTPKIEKIEVRKGRKVRRAKLYLV RSLRGAARIQIEIR" complement(10910..13513) /locus_tag="SE0917" complement(10910..13513) /locus_tag="SE0917" /codon_start=1 /transl_table=11 /product="conserved hypothetical protein" /protein_id="AA004514.1" /db_xref="GI:27315380" /translation="MIKNIWNKPIKRPFLIVLVCILSVLYGPELYQFITKGVIFSG SGDFQMPFQMHLYEHTFSGLGSDYFVKSLEYVSLYRLKPAPEFIATILYGM ILEQTNINVPDHSFTINQLIMAYVTVITPIFSFVLSVLRKPAPEFIATILYGM STVYTYNFTSWFSGNLLIMPLSLWAIRFFPKERKIGWIFAIATLFSNFPFSYE AIVGFYIFRPAIPHEKDIQVNRQKYLIVCATLLSVLSVLYGTYGVSFLDNWIL QNPENKIFTNLFETNYNIFADGFYITISFIAIALFCFLYOHYLYKFLAATWIL LIGFSQWFDSPAFNGFSLPQRWYFIALSTVLIALFIQHLSEISIKETYPFVILV IYGFIFALGERSVKMFVALIILIVLFIPIKYKSLTRTSMVLLVILFLAQQVLM NDSRKITTEPYOTKITINDSSVRSPLVKNKKIKYMQHSDTDLKRLDYFSYALNSPL IYHNGTSLVSGIFDDILKYDOTLQINMPVDKNSYRVLYNNRNANLMSLDVODLR HPDDLNMYPFKKELITDKQWIVHSNTINYPSAHITNKIYDARKLKSFLDREQM LKGVLNHSQANLDFPNLNLNNAQNINHANWIDSKHLVKQHGNGVTLNLPRI VKYQKMYIEMDVLLSFKDEHKVGVNEYSQERNLSYKVRVFSVPTWRAKASNQLN IKMSKVYRFKVGIGYENYOTLKASQOLQPVKVKERNGFTI IKKKHEHGYLVLP VYAEQHWAMANGPLKYOQNGINMTIIPVEGQTKILSYTPPYFYLLITVSCIGIL SILFTHVKKK" complement(13506..15779)	Qy	483	TGATAAGAACGCTTTTAAAGTGACGGTAAATGCGCAGAAATTGAATTTAAACAAAAACAGA	542
Ds		208935	TGATAAGAGTGCCTTTTAAAGTTACTGTATGCGCCATCAATTAGAAATTAACATAAACAGA	208876	
Qy		543	GTATGATTTACTATATCTTCTAGCTGAAATAAAAACCATGTTATGCAACGGGAACCAAT	602	
Ds		208875	ATACGATTTACTATATGTTTGTAGCTGAAATTCGTACCAACGTCATGCAACGTAACAAAT	208816	
Qy		603	TTTAAATCATGTATGGGGTATATAATAGTGAAGTAGAACAACAAATGTCGTAGATGTTTATAT	662	
Ds		208815	TCTCGATCAGGTATGGGGTATATAATAGTGAAGTAGAACAACGAATGTCGTATGTTTATAT	208756	
Qy		663	AAGATATTTACGAAACAAGTTAAACCATACGATCGTGAACAAATGATTCGAAACAGTTCCG	722	
Ds		208755	TCGTTATTTACGTAATAAATCAACCTTTTAAAGAAATAATCCATAGAAACAGTACG	208696	
Qy		723	TGCGTGTGGGTATGTCATACGATGACAAACG---TAAATTCGCGCAATAAATCGATTTAT	779	
Ds		208695	TGCGGTAGGGTATGTCATTCGATGATTAAAGCGCCAAAATAATAATATATAATGCGATGCTT	208636	
Qy		780	GTTTACCACGATGATTACGTTTGTGACGATATTTTGTGTTTGTGTTTAAATTTATTTTTC	839	
Ds		208635	ATTACGACGCTCATACCTTTCAGCAATATTAATCTTCTGCTTAAATTTATTTTCTTT	208576	
Qy		840	TTGAAAGATACACTGTCATATAGTACGCTTGTGATGATGCAAGACGAGCTCAACGCAATTT	899	
Ds		208575	TTAAAGATATCTTTACGAGTAGTGAAATTTGACGAGCTGAAAGAGTTTCAATGATATC	208516	
Qy		900	ATAATTTTATTTCAATCTAAGCCTGTTAAAGATATATCTGCATTAGACTTTGAATGCAATC	959	
Ds		208515	GCCAACTGTGTTCCATTTCTAAATCTTTAAAGTATATATCTGCAATGGAATTTAAATGCAATCC	208456	
Qy		960	TTAGGTAAATTTTCAAGAGATAATTTATTTATGATGAGCAATAATAATAATAATTTATTCGAGACA	1019	
Ds		208455	TTAGAAAAATTTTCAAGAAATTTGATTTATGATGATGATGAGGATGAGGATTTAATTTCAAAACA	208396	
Qy		1020	TCGAAATGATAACACAGTGAGAGTTGAACACAGGTTATGAACACCGTTATTTTTCACCGCGGTA	1079	
Ds		208395	TCAAATGATAATAACACTTCTGTTATGATAACAAAAATTGATTTTCAAACATCTCTGAACGTATA	208336	
Qy		1080	ATAAAAAACGCTATAAAGCGATTTGAATTTTAAATTTATTAAGAAACCAATTTACAACGCA	1139	
Ds		208335	CATATTTCAAAGAGACCGGATTAATTTACTTAGTAATAACTGAACTATACCTATACGTTTCGAA	208276	
Qy		1140	GATTTCAAAGGGGTATAGCTTTGTTTAAATTCATTCACCTAGAAAAATTTATGATAAACAATCGTAAAA	1199	
Ds		208275	GATTTTCTGGATACAGTGTATTTAGTCCCAATCTCTTTCAAANNTATGATTAATCTCGTTAAA	208216	
Qy		1200	TCATTTGATATCATTTGCGCTGGCATTTGGAGTATTTGCAACAAATTAATACTGCCACAATC	1259	
Ds		208215	TCATTTTATATAGTTGCACTTCTTTTGGATTAATTTGCAACCAATTTACTTCCCGCGGTG	208156	
Qy		1260	AGTTATGATTTTCAAACACAAATTTACTAAACCGCTTGTGAGTTTATCAAAATAAATGATTT	1319	
Ds		208155	AGTTATATCTTTCTTCAAAAATTTACTAAACCGATAGTTTACAATGTTCCAAATAAATGAAT	208096	
Qy		1320	GAGATTTGACGAGATGGTTTTTCAAATAAATTTGCAATTTAAATACAAATTTATGCAAGAAATA	1379	
Ds		208095	CAAAATAGAGAGATGGTTTTTCAAATAAATTTGAAATTTAACTTACAAATTTATGAGAAACA	208036	
Qy		1380	GATAATTTTAGCAAAATACGTTTAAATGAGATGATGAGCAAAATTTGAAGAAATCAATTTAATCAA	1439	
Ds		208035	GATAATTTTAAATTCATCTTTTAAATGAAATGATGATCAATAGAGAAATCTTTTAAATCAG	207976	
Qy		1440	CAAAGAACAATTTGTTGAAGATCGGTCAATGAATTTACGAACACCAATTTCAAATTTATCAA	1499	
Ds		207975	CAACGTCAAATTTGTCGAGGATGCTTCAACGAAATTTAAGAACGCCCACTGACGATTTATCAA	207916	
Qy		1500	GGTCATTTTAAATTTGATTCAGCGATGGGGAAAAAAGACCCGACGAGTTTGAAGAAATCG	1559	
Ds		207915	GGTCATCTAAATTTTAAATCCAAACGTTGGGGGAAAAAAGATCCAGCATGTTTGGAGAAATCT	207856	
Qy		1560	TTAAATATTTTCTATTGGAAGAAATGAATCGTATCATATAAATTTAGTCGAAGAAATTTACTTGA	1619	

Db	207855	TTGAATATTTCAATTTGAAGAGTGAATCGAATAACAAACCTTGTGGAAGACTTACTTTTA	207796
Qy	1620	TTGACTAAAGGAGATGTAATACACATTTCTTCTGAAGCGCAGACCGTGCATATTAATGAT	1679
Db	207795	CTTACCAAGATAGAGTCAATCATATGTTTGGGAATGTGAATAATGACGTAATAGC	207736
Qy	1680	GAATTCGCTCGCAATACACTCATATAAACAAATTCGATCTGATATCAATTTGATACG	1739
Db	207735	GAGATTTCAATCAAGTGTGAAGTCACTGCAACACCTACATCCAGATTATCTTTTGAACA	207676
Qy	1740	GATCTGACATCTAAAAATCTAGAAATTAATTAAGAACTTCATCAATTCGAACAAATTTT	1799
Db	207675	CATCTTGTCTAAGCCCTATCCAAATTAATAATTAACCGTCATCAGTTTGAACACTCTTA	207616
Qy	1800	TTAATCTTTATTTGATAATGCAATCAATATGATGTGAAGAATAAGAAAAATTAAGTTAAG	1859
Db	207615	CTCATATTTATGTAATGCAATGAAATACGACCTGAACATAGCAGATTAATTTGTT	207556
Qy	1860	ACAAGTTTAAAAATAAGCAAAAAATAATGAAATTAACAGATCATGGAAATTCGATATCCA	1919
Db	207555	ACTCAACTAAAAATAAATGATTAATGATGATGATTAATCTGATCATGGTATGGTATACCA	207496
Qy	1920	GAGGAAGATCAAGATTTTCATTTTTCATCGCTTTTATCGATGATTAATTCGTTCAAGA	1979
Db	207495	AAAGCTGACTTAGAATTTATCTTTGATAGATTTTATCGTGTAGATAAATCAGCTGCTCGT	207436
Qy	1980	AGTCAAGCGGTAAATGGACTCGGATTTATCTATTGCTCAAAAAATCATTAATTAACCGA	2039
Db	207435	AGTCAAGGAGCAATGGATTAGGACTATCAATAGCGGAAAAAATTTGTCGAATTAACCGT	207376
Qy	2040	GGATCGATTAATAATTAAGAGTGAATTAATAAGGAACAACGTTTAAATCATATTT	2096
Db	207375	GGTATGATTCAGGTAGAAAGTGAAGTCAACAAAGTACACGACTTTTCAAAATCAGTTT	207319
RESULT 10	AF269630	5230 bp DNA linear BCT 01-AUG-2000	
LOCUS	Staphylococcus epidermidis strain SRI clone step.1014f03 genomic		
DEFINITION	sequence.		
ACCESSION	AF269630		
VERSION	AF269630.1	GI:9623526	
KEYWORDS			
SOURCE	Staphylococcus epidermidis		
ORGANISM	Staphylococcus epidermidis		
REFERENCE	Bacteria; Firmicutes; Bacillales; Staphylococcus.		
AUTHORS	1 (bases 1 to 5230)		
	Kimmerly,W.J., Taylor,J.David., Nelsen,A.J., Godlevski,M.M.,		
	Rubino,M.A., Nelson,F.J., Rivers,P.R., Torruella-Miller,I.,		
	Listenbee,S., Ashanti,C., Althuller,G., Mamo,L., Shepherd,N.S.,		
	Fuchs,R., Fleming,T., Guan,X., Du,L., Cain,D.H., Miller,G.S. and		
	Furdon,P.J.		
TITLE	Transposon-mediated sequencing of the Staphylococcus epidermidis		
JOURNAL	genome		
REFERENCE	2 (bases 1 to 5230)		
AUTHORS	Taylor,J.David., Kimmerly,W.J., Nelsen,A.J., Godlevski,M.M.,		
	Rubino,M.A., Nelson,F.J., Rivers,P.R., Torruella-Miller,I.,		
	Listenbee,S., Ashanti,C., Althuller,G., Mamo,L., Shepherd,N.S.,		
	Fuchs,R., Fleming,T., Guan,X., Du,L., Cain,D.H., Miller,G.S. and		
	Furdon,P.J.		
TITLE	Direct Submission		
JOURNAL	Submitted (22-MAY-2000) Departments of Genomic Sciences and		
	Bioinformatics, Genetics Directorate, Glaxo Wellcome, Inc., 5 Moore		
	Drive, Research Triangle Park, North Carolina 27709-3398, USA		
FEATURES	Location/Qualifiers		
source	1. 5230		
	/organism="Staphylococcus epidermidis"		
	/mol_type="genomic DNA"		
	/strain="SRI"		
	/db_xref="taxon:1282"		
	/clone="step.1014f03"		

Db 723 ACTTTTCTGCTTAATATTTATTTCTTTTAAAGATATCTTTACGAAGTAGTGAATTTGA 782
Qy 872 TGATCGAGAACGAAGCTCAAGCGATATTAATAATTTATTTCTTAAGCCTGTTAAAGA 931
Db 783 CGAAGCTGAAGAGATTCAATGATATCGCCAACTGTTCCTAATCTTAATCTTTAAGTGA 842
Qy 932 TATATCTGCATTAGACTTGAATGCACTTTAGGTAAATTTTCAAGAGATAATTTATTTATGA 991
Db 843 TATATCTGCATTGGATTAAATGCACTTTAGAAAATTTTCAAGAAATCTGATCTATGA 902
Qy 992 TGAGCATTAATAATTTATTTGAGACATCGAATGATAACACAGTGAAGTTGAACGAG 1051
Db 903 TGATAACCGTAGAAAATTAATTCAAACATCAATGATAATACACTTGTCTATGATAACAA 962
Qy 1052 TTATGAACACCGCTTATTTTGACCGGTAAATAAAAAAGCTATAAAGCGCATTTGAATTT 1111
Db 963 AATTGATTTCAACATCTGGAAGTATATATATTCAGAGAAGCCAGGTATTAATTTACTT 1022
Qy 1112 AATTATTAAGAACCAATTTACACGCAAGATTTCAAGGGTATAGCTTTGTTAAATTCATTC 1171
Db 1023 AGTAATAACTGAACCTATACGTTTCAAGAGATTTTCTGGATACAGTGTATTTAGTCCATTC 1082
Qy 1172 ACTAGAAAATTTATGATAACATCGTAAATCATTTGATATATCATTTGCGTGGCATTTGGAGT 1231
Db 1083 TCTTCAAAAATTTATGATAATCTCGTTAAATCACTTTATATCGTTGCACTTGTCTTTGGATT 1142
Qy 1232 GATTGCAACAAATTTATAACTGCGCACAACTAGTTATGTTTTCACACAAATTTACTAAACC 1291
Db 1143 AATTGCAACCAATTTACTGCGGGGTGAGTTATATCTTTTCTTCCAAATTTACTAAACC 1202
Qy 1292 GCTTGTCTAGTTTATCAAAATAAATGATTGAGATTCGACGAGATGGTTTTCAAAAATAAAT 1351
Db 1203 GATAGTTACAAATGTTCCAAATAAATGAATCAAAATTAAGAGAGATGGTTTTCAAAAATAAAT 1262
Qy 1352 GCAATTAATACAAATTTATGAAGAAATAGATAAATTTAGCAATAGCTTTTAAATGAGATGAT 1411
Db 1263 TGAATTAATACAAATTTATGAAGAAACAGATAAATTTAATTTGATACCTTTTAAATGAAATGAT 1322
Qy 1412 GAGCAAAATTTGAAGATCATTTAAATCAACAAAGACAATTTGTTGAAGATGCGTCCACATGA 1471
Db 1323 GTATCAAAATAGAAAGATCTTTTAAATCAGCAAGCTCAATTTGCGAGGATGCTTTCACAGA 1382
Qy 1472 ATTCAGAACACCATTTACAAATTTATTTCAAGGTCAATTTAAATTTGATTCAGCGATGGGGAAA 1531
Db 1383 ATTAAGAACGCCACTGCGAGATTTATCAAGTCACTAAATTTAATCCACGTTGGGGAA 1442
Qy 1532 AAAAGACCCAGCAGTATTAGAGAAATCGTTAAATATTTCTATGTAAGAAATGAATCGTAT 1591
Db 1443 AAAAGATCCAGCAGTTTGGAGAAATCTTTGAAATATTTCAATGTAAGAAATGAATCGAAT 1502
Qy 1592 CATAAATTTAGTCGAAGATTAATTTGAATTTGACTAAAGAGATGTAAATGACATTTCTTC 1651
Db 1503 AACAAATTTGTCGAAGAACTACTTTTCTTACCAGAAATAGATGCAATCATATGTTTT 1562
Qy 1652 TGAAGCGCAGACCGTCATATTAATGATGAAATTCGCTCGCAATACACTCAATTAATAACA 1711
Db 1563 GGAATGTGAATGTAGAGCTAAATAGCGAGATTCATCACTGTAAGTCACTGCAACA 1622
Qy 1712 ATTCATCTCTGATTTATCAATTTGATGCGGATCTGACATCTAAATTTCTAGAAATTTAAAT 1771
Db 1623 CCTACATCCAGATTTATCTTTTGAACACATCTTGCTACTAAGCCCTATCCAAATTTAAAT 1682
Qy 1772 GAAACCTCATCAATTCGACAAATTTTAAATCTTTTATCTTTATGATGCAATGCAATATGA 1831
Db 1683 TAACCGTATCAGTTTGAACAACTCTTACTCATATTTATTTGATAATGCAATGAAATGCA 1742
Qy 1832 TGTGAAGAAATAAGAAAATTTAAAGTTAAAGCAAGGTTAAAAAATAAGCAAAAATAATTTGA 1891
Db 1743 CACTGACATAAGACATTTAAATTTGTTACTCAACTAAATAATTAATGATTTATGATGA 1802
Qy 1892 AATTACAGATCATGGAATTTGGTATTTCCAGAGGAAGATCAAGATTTCAATTTTGTATCGCTT 1951

Db 1803 TATTACTGATCATGTTGGGTATACCAAAAGCTGACTTAGAATTTATCTTTGTAGATTT 1862
Qy 1952 TTATCGAGTGAATAAATCTCGTTCAAGAAAGTCAAGCGGTAAATGACTCGGATTTATCTAT 2011
Db 1863 TTATCGTGTAGATAAATCACGTCGTCTGATGTCAGGAGGCAATGGATTAGACTATCAAT 1922
Qy 2012 TGCTCAAAAATCATTTCAATTTAAACGAGGATTCGATTTAAATTTAAAGTGAATTAATA 2071
Db 1923 AGCAAAAAAATTTGCAAACTTAACGGTGGTATGATTTCAAGTAGAAGAGTGAACATA 1982
Qy 2072 AGGACAAAGTTTAAATCATATTT 2096
Db 1983 GTACACGACTTTCAAAATCAGTTT 2007
RESULT 12
AF270244/c 3018 bp DNA linear BCT 01-AUG-2000
LOCUS Staphylococcus epidermidis strain SRI clone step.1056b10 genomic
DEFINITION sequence.
ACCESSION AF270244
VERSION AF270244.1 GI:9624154
KEYWORDS
SOURCE Staphylococcus epidermidis
ORGANISM Staphylococcus epidermidis
Bacteria; Firmicutes; Bacillales; Staphylococcus.
REFERENCE 1 (bases 1 to 3018)
AUTHORS Kimmerly, W.J., Taylor, J., David, J., Nelsen, A.J., Godlevski, M.M.,
Rubino, M.A., Nelsen, P.J., Rivers, P.R., Torruella-Miller, I.,
Listebbee, S., Ashanti, C., Altschuler, G., Mamo, L., Shepherd, N.S.,
Fuchs, R., Fleming, T., Guan, X., Du, L., Cain, D.H., Miller, G.S. and
Furdon, P.J.
TITLE Transposon-mediated sequencing of the Staphylococcus epidermidis
genome
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 3018)
AUTHORS Taylor, J., David, J., Kimmerly, W.J., Nelsen, A.J., Godlevski, M.M.,
Rubino, M.A., Nelsen, P.J., Rivers, P.R., Torruella-Miller, I.,
Listebbee, S., Ashanti, C., Altschuler, G., Mamo, L., Shepherd, N.S.,
Fuchs, R., Fleming, T., Guan, X., Du, L., Cain, D.H., Miller, G.S. and
Furdon, P.J.
TITLE Direct Submission
JOURNAL Departments of Genomic Sciences and
Bioinformatics, Genetics Directorate, Glaxo Wellcome, Inc., 5 Moore
Drive, Research Triangle Park, North Carolina 27709-3398, USA
FEATURES
source
1..3018
/organism="Staphylococcus epidermidis"
/mol_type="genomic DNA"
/strain="SRI"
/db_xref="taxon:1282"
/clone="step.1056b10"
ORIGIN
Query Match 47.3%; Score 1041.6; DB 1; Length 3018;
Best Local Similarity 71.5%; Pred. No. 1.1e-140;
Matches 1384; Conservative 0; Mismatches 549; Indels 3; Gaps 1;
Qy 164 ACAATGTGGACACAGAGTATGATGACAAAGCGTTTATAGATAAAGCGCTTAGCCATTACT 223
Db 3018 ATACTGTTGATATAGAGAAATGATGAAAGGTGGGTTTGGATAAAGCATTTATCTAAGCCCT 2959
Qy 224 ATGATTTAATCATATTAGATTTAATTTGTCGCTCAATTAATGGCTTAGAAAATTTCTCGCA 283
Db 2958 ATGATTTAATTTTATAGACTTTAATGCTTCCAAACATAAATGGTCTAGAAATTTAGAC 2899
Qy 284 AAATTTAGACAAACAATCTACCTATTTATTAATTTACGCGAAAAGTGTAGATGATG 343
Db 2898 AAATTCGTCAAAAACAACACTACTCCAAATTAATCATCTACTTACAAAAGCGAGACATATG 2839
Qy 344 ACAAGTTGCTGGGCTTGAATTCGGTGCAGAGGATTTATAGTTAAAGCGGTTTGTATG 403
Db 2838 ATAAAGTAGCTGGGTTGGACTATGGGCGAGATGACTACTAATTTGTAACCCCTTTGTATAG 2779

```
Qy 404 AAGAACTTTTACCAAGAAATTCGTCAAAATTTTACGTCTGTCAGCCCAAAAGGATATTATCG 463
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2778 AAGAAATTCCTCGCAAGAAATAGACGGGTATTGCGCAGACAGCCAGATAAAGATGTTTTAG 2719
Qy 464 ATGTCAACGGTATTACAAATTCGATAAGACGGTTTTTAAAGTGACGGTAAATGGCCGCAAAA 523
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2718 ATATCAATGGTATTATCAATTCGATAAGATGCTTTTAAAGTTACTGTGTTAAATGGCCATCAAT 2659
Qy 524 TTGAATTAACAAAACAGAGATGATGATTTTAAATCATGTATCGGGTTATAATAGTCAAGTAGAACA 643
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2658 TAGAATTTAACTTAAACAGAAATACGATTTATATATGTTTTAGCTGAAATTCGTAACCAACG 2599
Qy 584 TTATGCAACGGGAACAAATTTTAAATCATGTATCGGGTTATAATAGTCAAGTAGAACA 643
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2598 TCATGCGCGGTGAACAAATTCGATCAAGTATCGGGGTATAATAGTGAAGTAGAACA 2539
Qy 644 ATGTGCTAGATGTTTTATATAGATATTTCGAACAAGTTTAAACCAATAGATCGTGACA 703
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2538 ATGTGCTGTGATGTTTACATTCGTTGTTTACGTAATAAATCAAACTTTTAAATAAGAAA 2479
Qy 704 AAATGATTTGAAACAGTTTCGTGGCGTTCGGTATGTATGATACGATGACAAACG---TAAAT 760
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2478 AATCCATAGAAACAGTACGTGGCGTAGGTATGTATGATTCGAATTAAGCGCCAAAAT 2419
Qy 761 GCGCAATACTGGATTTATGTTACCAAGATGATTACGTTTGTCAAGATATTTTTTGTGTTG 820
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2418 AAATATAAATGGATGCTTATTACGACGCTCATTTACCTTCACGACATATTTACTTTCTG 2359
Qy 821 TTTAATTTATTTTTTTCTGAAAGATACACTGCATAATAGTGAGCTTGATGATGACAGA 880
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2358 CTTAATTTATTTTCTTTTAAAGATACCTTTACGAAAGTAGTGAATTTTACGCAAGCTGA 2299
Qy 881 ACGAAGCTCAAGCGATATTAATAATTTATTTCTAAGCTGTAAAGATATATCTGC 940
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2298 AAGAAATTCAAATGATTCGCAATCTGTTCCATCTTAAATCTTTAAGTATATCTGC 2239
Qy 941 ATTAGACTGAATGCATCTTTAGTAAATTTTCAAGAGATAATTTATTTATGATGAGCATAA 1000
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2238 ATTGAAATTAATGCATCTTTAGAAAATTTTCAAGAAATCTGATCTATGATGATAACGG 2179
Qy 1001 TAATAAATTTTGAAGATGAAATGATACACAGTGAGAGTTGAACCGGTTATGAAACA 1060
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2178 TAGAAAGTTAATTCAAACATCAATATGATAATACACTTGTCTTATGATAACAAAATTTGAT 2119
Qy 1061 CCGTATTTTTCGCGCGTAATAAAAACGCTATAAAGGCATTTGAATTTTAAATTTAA 1120
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2118 CAACATCTCTGAACGTAATACATTTCAAGAGCCGCGTATTAAATTTACTTTAGTAATAAC 2059
Qy 1121 AGAACCAATTACAAACGAAGTTTCAAAGGGTATAGCTTGTAAATTCATTCACCTAGAAAA 1180
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2058 TGAACCTATAGCTTCGAAAGATTTTCTGGATACAGTGTATTAGTCCATCTCTTCAAAA 1999
Qy 1181 TTATGATAACATCGTAAATCATTTGATATATCATTTGCGCTGGCATTTGGAGTATGCAAC 1240
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1998 TTATGATAATCTCGTTAAATCACTTTATATCGTTGCACTTGTCTTTGGATTAATTTGCAAC 1939
Qy 1241 AATTATACCTGCCAATCAGTTATGATTTTCAACACAAATTAATTAACCGCTTGTGAC 1300
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1938 CATTTATCTGCGCGGTGAGTTATATCTTTCTTCGCAATTAATTAACCGATAGTTAC 1879
Qy 1301 TTTATCAAAATAAATGATTTGAGATTTGACGAGATGTTTTTCAAAATAAATTTGCAATTTAA 1360
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1878 AATGTCCAAATAAATGAATCAAAATTAGAAGATGTTTTTCAAAATAAATTTGATTAAC 1819
Qy 1361 TACAATTTAGAAATATAGATTAATTTAGCAAAATACGTTTAAATGAGATGATGACCAAT 1420
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1818 TACAATTTAGAAATACAGATAATTTAATTTGATCTTTTAAATGAAATGATGATCAAT 1759
Qy 1421 TGAAGATCATTTAATCAACAAAGACAAATTTGTTGAAGATGCGTCACATGAATTTACGAAC 1480
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1758 AAGAAATCTTTTAAATCAGCAACGTCAATTTTGTGAGGATGCTTCACACCAATTTAAGAAC 1699
```

```
Qy 1481 ACCATTACAAATTTATTCAAGGTCAATTTAAATTTTGAATTCAGCGATGGGAAAAAAGACCC 1540
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1698 GCCACTGCAGATTTATTCAAGGTCACTTAAATTTAAATCCAAAGTTGGGGGAAAAAAGATCC 1639
Qy 1541 AGCAGTATTAGAAGATCGTTAAATATTCTTATTGAAGAAATGAATCGTATCATAAAT 1600
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1638 AGCAGTTTGGGAAGAATCTTTTGAATATTTCATTTGAAGAAAGTGAATCGAATAACAAACT 1579
Qy 1601 AGTCGAAGATTTACTTGAATTTGACTTAAAGGAGATGTAATGACATTTCTTCTGAAGCGCA 1660
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1578 TGTCGAAGAACTACTTTTACTTACAAAGATAGAGTCAATCATATAATGTTTGGAAATGTGA 1519
Qy 1661 GACCGTCATATTAAATGATGAAATTCGCTCGCAATACACTCATTTAAACAATTTGCATCC 1720
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1518 AAATGTAGACGTAATAGCGAGATTCAATCACGTGTCAAGTCACTGCAACACCTACATCC 1459
Qy 1721 TGATTTCAATTTTGATACGGATCTGACATCTAAATAATCTAGAAATTTAAATAGAAACCTCA 1780
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1458 AGATTATACTTTTGAACACATCTTGCTACTAAGCCCTATCCAATTTAAATAATTAACCGTCA 1399
Qy 1781 TCAATTCGAACAATTTATTTTAAATCTTTTATGATAATGCAATCAAAATATGATGTGAAGAA 1840
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1398 TCAGTTTGAACAACACTCTTACTCATATTTATGATAATGCAATGAAATACGACACTGAACA 1339
Qy 1841 TAAGAAAATTTAAAGTTTAAAGCAAGTTTAAAAAAATAAGCAAAAAATTAATTTGAAATTTACAGA 1900
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1338 TAAGCAATTTAAATTTGTTACTCAACTCAACTAAAAAATAAATGATTATGATTGATTAATCTGA 1279
Qy 1901 TCATGAATTTGGTATTTCCAGAGGAAGATCAAGATTTTCATTTTTTGTATGCTGCTTTTATCGAGT 1960
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1278 TCATGTTATGGGTATACCAAAAGCTGACTTTAGAAATTTATCTTTGATAGATTTTATCGTGT 1219
Qy 1961 GGATAAATCTCGTTCAAGAGAGTCAAGCGGTATGACTCGGATTAATCTATCTATCTGCTCAAAA 2020
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1218 AGATAAATCAGTCTCGTAGTCAAGAGCAATGAGTATAGGACTATCAATAGCAGAAAA 1159
Qy 2021 AATCATTTCAATTTAAACGGAGATCGATTAAAAATTTAAAGTGAATTAATTAAGGAAACAAC 2080
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1158 AATTGTCAACTTTAACGGTGTATGATTCAAGTAGAAGTGAAGTCAAAAAGTACACGAC 1099
Qy 2081 GTTTAAATCATATTTT 2096
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1098 TTTCAAAATCAGTTTT 1083

RESULT 13
AX14562/c
LOCUS 3018 bp DNA linear PAT 31-MAY-2001
DEFINITION Sequence 4284 from Patent WO0134809.
ACCESSION AX145562
VERSION AX145562.1 GI:14284127
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM artificial sequences.
REFERENCE 1
AUTHORS Kimmerly, W.J.
TITLE Staphylococcus epidermidis nucleic acids and proteins
JOURNAL Patent: WO 0134809-A 4284 17-MAY-2001;
GLAXO GROUP LIMITED (GB)
FEATURES
source
1. .3018
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="synthetic nucleic acid sequence"
ORIGIN
Query Match 47.3%; Score 1041.6; DB 6; Length 3018;
Best Local Similarity 71.5%; Pred. No. 1.1e-140;
Matches 1384; Conservative 0; Mismatches 549; Indels 3; Gaps 1;
Qy 164 ACAATGTGACACAGAGTATGATGGACAAGACGGTTTGTAGATAAAGCGCTTAGCCATTACT 223
```

Db 3018 ATACTGTTGATATAGAGATGATGAAAGGTGGGTTTGATAAAGCAATATCTAAGCCCT 2959
Qy 224 ATGATTTAATCATATATAGATTTAAAGTTGCGGTCAATTAATGGCTTAGAAATTTCTCGCA 283
Db 2958 ATGATTTATATATTTAGACCTTTAATGCTTCCAAACATAAATGGCTAGAAATTTGTAGAC 2899
Qy 284 AAATAGACAACAACAATCTACACCTATCATTTAATTAACGCGAAAAGTGATCGTATG 343
Db 2898 AAATTCGTCAAAACAACTACTCCAAATATCATCTACTTGCACAAAAGCGAGACATATG 2839
Qy 344 ACAAGTTGCTGGGCTTGATTTAGGTGCGAGCAGATATATAGTTAAGCGGTTTGATATG 403
Db 2838 ATAAAGTAGCTGGGTTGGACTATGGGCGAGATGACTACATTTGAAAACCCCTTTGTATAG 2779
Qy 404 AAGAACTTTTACCAAGAAATTCGTGCAATTTTACGTGCTGAGCCACAAAGAGATATATCG 463
Db 2778 AAGAATTCGTCGAAGAATAAGCGGTATTCGGCAGACAGCATAAAGATGTTTTAG 2719
Qy 464 ATGTCACCGGTATTACAAATTTGATAAGAACGCTTTTAAAGTGACGGTAAATGGCGCAGAA 523
Db 2718 ATATCAATGGTATTATCATTTGATAAAGATGCCCTTTAAAGTTACTGTTAATGSCCATCAAT 2659
Qy 524 TTGAATTAACAAAACAGAGATGATTTACTATATCTTCTAGCTGAAAATAAAAACCATG 583
Db 2658 TAGAAATTAACATAACAGAAATACGATTTATTATATGTTTGTAGCTGAAAATTCGTAACCA 2599
Qy 584 TTATCCAAACGGGAAACAAATTTTAAATCATGATGAGGTTTATAATAGTGAAGTAGAACA 643
Db 2598 TATGACAGCGTGAACAAATTCGATACGATGAGGTTTATAATAGTGAAGTAGAACA 2539
Qy 644 ATGTCGTAGATGTTTATATAAGATTTTACGAAACAAAGTTAAACCATACGATCGTGACA 703
Db 2538 ATGTCGTGATGTTTACATTCGTTATTTACGTAATAAATCAACCTTTTAAATGAAGAA 2479
Qy 704 AAATGATTGAACAGTTCTGGCGGTTGGTATGATGATGATGATGATGATGATGATGATGAT 760
Db 2478 AATCCATAGAAACAGTACGTCGCGTAGGTTATGATTCGATGATGATGATGATGATGAT 2419
Qy 761 GCGCAATACGATGATTTGTTACCAAGATGATGATGATGATGATGATGATGATGATGATG 820
Db 2418 AAAATATAAATGGATGCTTTATACGACGCTCAATACCTTCAGCAATATATCTTTCTG 2359
Qy 821 TTTAATATTTATTTTCTTGAAGATACACTGCATATAGTAGTACGCTTGATGATGACAGA 880
Db 2358 CTTAATATTTATTTCTTTTAAAGATACCTTTACGAGTAGTGAAATGACGAGCTGA 2299
Qy 881 ACGAAGCTCAACGGATATTAATAATTTATTTCAATTCAGCTGTTAAAGATATATCTGC 940
Db 2298 AAGAAGTTCAAAATGATATCGCCAATCTGTTCCAAATCTTAAATCTTTAAGTGATATCTGC 2239
Qy 941 ATTAGACTTGAATGATCTTTAGTAAATTTTCAAGAGATATTTATTTATGATGACATAA 1000
Db 2238 ATTGATTTAAATGATCTTTAGAAATTTTCAAGAAATCTGATCTATGATGATAACGG 2179
Qy 1001 TAATAAATTTATGAGACATCAATGATAACACAGTGAGAGTTGAACCAAGGTTATGAACA 1060
Db 2178 TAGAAGTTAATTCAAACATCAATGATATATACCTTGCTTATGATAACAAATTTGATTT 2119
Qy 1061 CGGTATTTTGAACGGTAAATAAAGAGCTATTAAGGCAATGAAATTTAATTTATTTAA 1120
Db 2118 CAACATCTGAACGATATACATATTTCAAGAGACCAACGGTATTAATTTACTTAGTAATA 2059
Qy 1121 AGAACCAATTACAAACGAGATTTCAAGGGTATAGCTTGTAAATCTTCACTAGAGAAA 1180
Db 2058 TGAACCTTACGTTGCAAGAGATTTTCTGGGATACAGTGTATTGCTTCTTTCAAAA 1999
Qy 1181 TTATGATAACATCGTAAATCAATTTGATATCATTTGCGCTGGCATTTGGAGTATGCAAC 1240
Db 1998 TTATGATAATCTCGTTAATACATTTATATCGTTGCACTTGTGTTTGAATTAATTTGCAAC 1939
Qy 1241 AATTATACTGCCAAATCAGTTATGATTTTCAACACAAATTAATAACCCGCTTGTGAG 1300

Db 1938 CATTATTACTCCGGCGTGAGTTATATCTTTCTTCGCAAAATTACTAAACCGATGTTAC 1879
Qy 1301 TTTTATCAAAATAAATGATTGAGATTCGACGAGATGGTTTTCAAAATAAATTTGCAATTTAA 1360
Db 1878 AATGTCATTAATAATGATCAATTTAGNAGAGATGGTTTTCAAAATAAATTTGATTAAC 1819
Qy 1361 TACAAATTTAGAGAAATAGATAAATTTTAGCAAAATACGTTTTAATGAGATGATGAGCAAA 1420
Db 1818 TACAAATTTAGAGAAACAGATAAATTTAATTTGATCTTTTAAATGAAATGATGATCAAT 1759
Qy 1421 TGAAGATCAATTTAATCAACAAAGACAAATTTGTTGAAGATCGTCAATGAATTAACGAC 1480
Db 1758 AGAAGATCTTTTAAATCAGCAACGTCATTTGTGAGGATGCTTCAACAGCAATTAAGAAC 1699
Qy 1481 ACCATTACAAATTTAATCAAGGTCAATTTAAATTTTCAGATTCAGCGATGGGGAAAAAGACCC 1540
Db 1698 GCCACTGCAGATTAATCAAGGTCAATTTAATTTAATCCAGCTTGGGGGAAAAAAGATCC 1639
Qy 1541 AGCAGTATTAGAGAAATCGTTTAAATTTTCTTATTTGAAGAAATGAATCGTATCATATAAAT 1600
Db 1638 AGCAGTTTGGAGAAATCTTTTGAATATTTCAATTTGAAGAGTGAATCGAATAACAAACT 1579
Qy 1601 AGTCGAAGAAATTTACTTGAATTTGACTTAAAGGAGATGTTAAATGACATTTCTTCTGAAGCC 1660
Db 1578 TGTGGAAGAACTCTTTTACTTACCAAGATAGAGTCAATCAATGTTTGGAAATGTGA 1519
Qy 1661 GACCGTGATATTAATGATGAAATTCGCTCGCAATACACTCAATTTAAACAAATTTGCATCC 1720
Db 1518 AAATGTAGACGTAATAGCGAGATTCATCACTGCTGAGTCACTGCAACACCTACATCC 1459
Qy 1721 TGATTATCAATTTGATACGGATCTGACATCTAATAAATCTAGAAATTTAAATTTGAACCTCA 1780
Db 1458 AGATTATCTTTTGAACACATCTTGTCTACTTAAGCTTATCCAAATTTAAATAATTAACCGTCA 1399
Qy 1781 TCAATTCGAACAAATTTTAAATCTTTTATTTGATTAATGCAATCAATGATGATGAAGAA 1840
Db 1398 TCAGTTTGAACAACTCTTACTCATATTTTATTTGATTAATGCAATGAAATACGACACTGA 1339
Qy 1841 TAAGAAATTTAAAGTTTAAGCAAGTTTAAATAAAGCAAAATAATTTGAAATTTACAGA 1900
Db 1338 TAAGCAATTAATAATTTGTTACTCAACTAATAAATAATGAATGATGATGATTAATCTGA 1279
Qy 1901 TCATGGAATTCGTTATTCAGAGGAGATCAAGATTTTCATTTTTCATCGCTTTTATCGAGT 1960
Db 1278 TCATGTTATGGTATATACCAAAAGCTGACTTGAATTTTATCTTTGATAGATTTTATCGTGT 1219
Qy 1961 GGATAAATCTCGTTCAAGAGTCAAGCGGTAAGGCTGAGTCAATTTCTATTTGCTCAAAA 2020
Db 1218 AGATAAATTCACGTCGTCGTAGTCAAGGAGGCAATGGGATTAGGACTATCAATAGCAGAAA 1159
Qy 2021 AATCATTTCAATTAACCGGAGATCGATTTAAATAATTTAAAGTCAAAATTAATAAGGAACAAC 2080
Db 1158 AATTGTCAACTTAACGGTGTATGATTTCAAGTAGAAGTGAAGTGAAGTGAAGTGAAGTGA 1099
Qy 2081 GTTTAAATCATATTTT 2096
Db 1098 TTTCAAAATCAGTTT 1083

RESULT 14
AF269746/c

LOCUS

DEFINITION
Staphylococcus epidermidis strain SRI clone step.1021e09 genomicACCESSION
AF269746VERSION
AF269746.1KEYWORDS
GI:9623642SOURCE
Staphylococcus epidermidisORGANISM
Staphylococcus epidermidisREFERENCE
1. (bases 1 to 3099)AUTHORS
Kimmerly,W.J., Taylor,J.David., Nelsen,A.J., Godlevski,M.M.,
Rubino,M.A., Nelson,F.J., Rivers,P.R., Tortuella-Miller,I.,

Listenbee, S., Ashanti, C., Altschuller, G., Mamo, L., Shepherd, N.S.,
 Fuchs, R., Fleming, T., Guan, X., Du, L., Cain, D.H., Miller, G.S. and
 Furdon, P.J.
 Transposon-mediated sequencing of the *Staphylococcus epidermidis*
 genome
 Unpublished
 2 (bases 1 to 3099)
 Taylor, J. David., Kimmerly, W.J., Nelsen, A.J., Godlevski, M.M.,
 Rubino, M.A., Nelson, F.J., Rivers, P.R., Torruella-Miller, I.,
 Listenbee, S., Ashanti, C., Altschuller, G., Mamo, L., Shepherd, N.S.,
 Fuchs, R., Fleming, T., Guan, X., Du, L., Cain, D.H., Miller, G.S. and
 Furdon, P.J.
 Direct Submission
 Submitted (22-MAY-2000) Departments of Genomic Sciences and
 Bioinformatics, Genetics Directorate, Glaxo Wellcome, Inc., 5 Moore
 Drive, Research Triangle Park, North Carolina 27709-3398, USA
 Location/Qualifiers
 source
 1. 3099
 /organism="Staphylococcus epidermidis"
 /mol_type="genomic DNA"
 /strain="SRI"
 /db_xref="taxon:1282"
 /clone="step.1021e09"
 ORIGIN
 Query Match 44.5%; Score 978.8; DB 1; Length 3099;
 Best Local Similarity 70.9%; Pred. No. 1.1e-131;
 Matches 1314; Conservative 0; Mismatches 537; Indels 3; Gaps 1;
 3 ATTTAAAGCACTATTGTATAGAAAATACAAAATTTAAATATATATTACCTTTATAGA 62
 1854 AGTTTCATACAAAATCATTTTACAAATTTTGAATTACATCTTTAGAAATACGAGTTAATGT 1795
 63 AAAAGTCGTAATATGAGGTGACAAATGACGCAAAATTTTAAATAGTAGAAGATGAACAAA 122
 1794 TTAAGATGTGCGGTGAGGTGCAATATGACAAATATTTTAAATGTAGAAATGAGCAAAA 1735
 123 CTTAGCAAGATTTCTTCAATTTGGAATCAGACATGAAATTTACAAATGTTGACACAGTA 182
 1734 TCTTGCTAGATTTATAGAGCTTGAGTTAATCATGAAATTTATCTGTTGATATAGAA 1675
 183 TGATGCAAGACGGTTTATAGTAAAGCGCTTAGCCATTACTATGATTTAATCATATTAGA 242
 1674 TGATGGAAGGTGGTTTGATTAAGCATTTATCTAAGCCCTATGATTTATATTATTAGA 1615
 243 TTTAATGTTCCGTCATTAATGCTTAGAAATTTGCGCAAAATTTAGACAAACAATC 302
 1614 CTTAATGCTTCCAAACATAAATGGTCTAGAAATTTGTAGACAAATTCGTCAAAAACAAAC 1555
 303 TACACCTATCATTTAATTACAGCAAAAGTGATAGTATGACAAAGTTGCTGGGCTTGA 362
 1554 TACTCCAATTTATCATCTACTGCAAAACCGAGACAAATGATTAAGTAGCTGGTTGGA 1495
 363 TTACGGTGCAGACGATTTATATAGTTAAGCGGTTTGTATTTGAAGAATCTTTTACAAAGAT 422
 1494 CTATGGGCAGATGACTACATTTGTAATGTAACCCCTTTGTATATAGAAATTTGCTCGCAAGAT 1435
 423 TCGTGCAATTTTACGTCGACGCCAACAAAGGATATTTATCGATGTCACGGTATTAACAAT 482
 1434 AAGAGCGGTTATGCGCAGACAGCCAGATAAAGATGTTTTAGATATCAATGGTATTATCAT 1375
 483 TGATAAGACGCTTTTAAAGTGAAGTAAATGCGCAGCAATTTGAATTAACAAACACAGA 542
 1374 TGATAAGATGCTTTTAAAGTGAAGTAAATGCGCAGCAATTTGAATTAACAAACACAGA 1315
 543 GTATGATTTACTATATCTTCTAGCTGAAAATAAAAAACCACTGTTATGCAACGGGAACAAAT 602
 1314 ATACGATTTTATATATGTTTATGCTGAAAATCGTAACCACTGCTATGACGGCTGAACAAAT 1255
 603 TTTAATCATGATGGGTTTATTAATAGTGAAGTAGAAGCAAAATGCTGATGATGTTTATAT 662
 1254 TCTCGATCAGTATGGGGTAAATATGTAAGTAGAAGCAAAATGCTGATGTTTATAT 1195

QY 663 AAGATATTTACGAAACAGTTTAAACCATACGATCGTGACAAATGATTGAAACAGTTCG 722
 Db 1194 TCGTTATTTACGTAATAAATCAACCTTTTAAATAAGAAAAATCCATAGAAACAGTACG 1135
 QY 723 TGGCGTTGGGTATGTGATACGATGACAAACG---TAAATTTGCGCAATTAACCTGGATTAT 779
 Db 1134 TGGCGTAGGTATGTGATTCGATGATTAAGCGCAAAATTAATAATATATAATGGATGCTT 1075
 QY 780 GTTACCACGATGATTAAGTTGTCAGATATTTTGTGTTTGTGTTAATATTATTTTTC 839
 Db 1074 ATTACACGCTCATTTACCTTCACGACATATTAATTTCTGCTTAATATTATTTCTTT 1015
 QY 840 TTGAAAGATACACTGCATAATAGTACGTTGATGATGACAAACGAGCTCAACGATATT 899
 Db 1014 TTAAGATATCTTTACGAGTAGTGAATTTGACAAAGCTGAAAGAGTTTCAATATGATATC 955
 QY 900 AATAATTTATTTTCAATCTAAGCCTGTTAAAGATATATCTGCATTAGACTTTGAATGCATCT 959
 Db 954 GCCAATCTGTTCCATTTCTAAATCTTTAAGTGATATATCTGCATTGGATTTAATGATCC 895
 QY 960 TTAGGTAAATTTCAAGAGATATTTTATGATGACGATATAATTAATAATTTATTTGAGACA 1019
 Db 894 TTAGAAAAATTTTCAAGAAATACTGATCTATGATGATAACGGTAGAAGTTAATTTCAACA 835
 QY 1020 TCGAATGATAACACAGTGAGAGTTGAACCCAGGTTATGAACACCGTTATTTTGACCCGTA 1079
 Db 834 TCNAATGATAATACACTCTGCTTATGATAACAAATTTGATTTCAACATCTCTGAACTGATA 775
 QY 1080 ATAAAAAAGCGCTATTAAGGCGATTTGAATATTTAAATTTTAAAGAACCAATTTACAGCGAA 1139
 Db 774 CATATTTCAAGAAAGCCACGCTATTAATTTACTTAGTAATACTGAACCTATACGTTGMAA 715
 QY 1140 GATTTCAAGGGTATAGCTTTGTTAATTTCAATCTACAGAGAAATTTATGATTAACATCGTAA 1199
 Db 714 GATTTTCTCGATACAGTGTATAGTCCATTTCTCTTCAAAATTTATGATTAATCTCGTTAA 655
 QY 1200 TCATTGTATATCATTCGCTGCGCTTTGGAGTGTGCAACCAATTAATACTGCACAATC 1259
 Db 654 TCATTTTATCGTTCGACTTTCCTTTGGATTAATTTGCAACCAATTTACTGCGCGGTG 595
 QY 1260 AGTTATGATTTTCAACACAAATTTACTAAACCGCTTTGTGAGTTTATCAATTAATGATTT 1319
 Db 594 AGTTATATCTTTCTTCGCAATTTACTAAACCGATAGTTTACAATGCTCAATTAATGAT 535
 QY 1320 GAGATTCGACGAGATGGTTTTCAAATAAATTTGCAATTTAAATACAAATTTATGAAGAAAT 1379
 Db 534 CAAATTTAGAGAGATGGTTTTCAAATAAATTTGAAATTAACCTACAAATTTATGAAGAAAC 475
 QY 1380 GATAATTTAGCAATACGTTTAAATGATGATGAGCGCAATTTGAAGCAATTTAATCA 1439
 Db 474 GATAATTTAATTTGATATCTTTAATGAATGATGATCAATTAAGAGAAATCTTTTAAATCAG 415
 QY 1440 CAAAGCAATTTGTTGAAGATGCTGATCAATGAATTTACGAACCACTTACAAATTTATTC 1499
 Db 414 CAAAGCAATTTGTTGAGGATGCTTTCACACGAAATTAAGAACGCCACTGCGAGATTTATCA 355
 QY 1500 GGTCAATTTAAATTTGATTCAGCGATGGGAAATAAGAACCCAGCATTTAGAGAAATCG 1559
 Db 354 GGTCAATTTAAATTTAATTCACCGTTGGGGGAAAAAGATCCAGCATTTTGGAGAAATCT 295
 QY 1560 TTAATATTTCTTATTTGAAGAAATGATCGTATCATAAATTTAGTCCGAAGAAATTTACTTGA 1619
 Db 294 TTGAATTTTCAATTTGAAGAGTGAATCGAATAACAAATCTTGTGCAAGAACTACTTTTA 235
 QY 1620 TTGACTTAAAGGAGATGTAATGACATTTCTTCTGAGCGCAGACCGTGCATTTAATGAT 1679
 Db 234 CTTTACCAAGATAGAGTCAATCATTAATGTTTGGAAATGTAAGAAATGTAAGCAATATAGC 175
 QY 1680 GAAATTCGCTCGGAAATACACTCATTAATAACAAATTTGATCTCTGATTTATCAATTTGATACG 1739
 Db 174 GAGATTTCAATCGTGTGAAGTCACTGCAACACCTACATCCAGATTTATATCTTTTGAACA 115
 QY 1740 GATCTGACATCTAAAAATCTAGAAATTTAAATGAAACCTTCATCAATTTGCAACAATTTT 1799

Db 114 CATCTGCTACTAAGCCTATCCAAATTAATAAATTAACCGTCATCAGTTTGAACAACTCTTA 55
Qy 1800 TTAATCTTTATTGATAATCAATCAATATGATGTGCAAGATAAGAAATTAATA 1853
Db 54 CTCATATTTATTGATAATGCAATGAATGAATACGACACTGAACATAGCACATTAATA 1
RESULT 15
AX145064/c 3099 bp DNA linear PAT 31-MAY-2001
LOCUS Sequence 3786 from Patent WO0134809.
ACCESSION AX145064
VERSION AX145064.1 GI:14283629
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM artificial sequences.
1
REFERENCE Kimmerly, W.J.
AUTHORS Staphylococcus epidermidis nucleic acids and proteins
TITLE Patent: WO 0134809-A 3786 17-MAY-2001;
JOURNAL GLAXO GROUP LIMITED (GB)
FEATURES Location/Qualifiers
source 1..3099
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="synthetic nucleic acid sequence"
ORIGIN
Query Match 44.5%; Score 978.8; DB 6; Length 3099;
Best Local Similarity 70.9%; Pred. No. 1.le-131;
Matches 1314; Conservative 0; Mismatches 537; Indels 3; Gaps 1;
Qy 3 ATTTAAAGCAACTATTCTGATAGAAAAATACAAAATTTAAATATATACCTTTATAGA 62
Db 1854 AGTTTCATACAAATCATTTTCAATTTTGAATTTACATCTTTAGAAATACGAGTTAATGT 1795
Qy 63 AAAAGTCGAATATGAGGTGTCAAATGACGCAATTTTAAATAGTAGAGATGACAAA 122
Db 1794 TTAAGATGAGCGTGAGGTGCAATATGACAAATTTTAAATGAGAGATGACAAA 1735
Qy 123 CTTAGCAAGATTTCTTGAATTTGGAATCACAATGAAATTTACAATTTGAGGACACAGATG 182
Db 1734 TCTGCTAGATTTATAGACCTTGAGTTAACTCATGAATTTATCTGTTGATAGAGAA 1675
Qy 183 TGATGACAAAGCGGTTTATAGATAAAGCGCTTAGCCATTTACTATGATTTAATCATATTAGA 242
Db 1674 TGATGAAAGGTGGGTTTGGATAAAGCAATTTCTAAGCCCTATGATTTATATATTTAGA 1615
Qy 243 TTTAATGTTCCGCTCAATTAATGCGTTAGAAATTTGTCGCAAAATTTAGACAAACATC 302
Db 1614 CTTAATGCTTCCAAACATAAATGGTCTAGAAATTTGTAGACAAATTTGTCGCAAAACAA 1555
Qy 303 TACACCTATCATTTAATTTACAGCGAAAGTGATGATGATGATGATGATGATGATGATGAT 362
Db 1554 TACTCAATTTATCACTTACTGCAAAAGCGAGACATATGATTAAGTAGCTGGGTTGA 1495
Qy 363 TTACGCTGACAGCATTTATATAGTTAAGCGGTTTGATTTGATTTGATTTGATTTGATTTG 422
Db 1494 CTATGGGGCAGATGACTACATTTGTAAGCCCTTTGATATAGAAATTTGCTGCAAGAT 1435
Qy 423 TCGTCAATTTTACGTCGTCAGCCCAAAAGGATTTATCGATGTCAGCGGTATTACAT 482
Db 1434 AAGACGGGTATTGCGCAGACACCCAGATAAAGATGTTTGTAGATATCAATGGTATTATCAT 1375
Qy 483 TGATAAGAACCGCTTTTAAAGTGACGTTAAATGCGCAGAAATTTGAATTAACAAAACAGA 542
Db 1374 TGATAAGAGATGCTTTTAAAGTTTACTGTTAATGGCCATCAATTTAGAAATTTACTTAAAC 1315
Qy 543 GTATGATTTTACTATATCTTCTAGCTGAAATTAATAAACCATGTTATGCAACGGGAACAAT 602
Db 114 ATACGATTTATATATGTTTTAGCTGAAAAATCGTAACCAACGCTCATGACGCGTAACAAAT 1255
Qy 603 TTTAAATCATCTATGGGCTTATAATAGTGAAGTAGAAAACAAATGTCGTAGATGTTTATAT 662
Db 1254 TCTCGATCACGATATGGGGGTATATAGTGAAGTAGAAAACGAAATGTCGTGATGTTTACAT 1195
Qy 663 AAGATATTTACGAAACAAAGTTTAAACCATACGATCGTGACAAAATGATTTGAAACAGATTG 722
Db 1194 TCGTTATTTTACGTAATAAATCAAACTTTTAAATAAGAAAAATCCATAGAAAACAGTAGC 1135
Qy 723 TGGCGTTGGGTATGATGATACGATGACAAAACG---TAAATTCGCGCAATAACTGGATTAAT 779
Db 1134 TGGCGTAGGGTATGATGATTCGATTAAGCGCCCAAAAATTAATAATTAATAATGGATGCTT 1075
Qy 780 GTTACACGATGATTACGTTTGTACAGATATTTTGTGTTGTTTAAATTTATTTTTC 839
Db 1074 ATTACGAGCTCATTTACCTTCAGCAATATATTCTTTCTGCTTAATTTATTTTCTTT 1015
Qy 840 TTGAAAGATACACTGTCATAATAGTAGCTTGATGATGACAGAACGAGCTCAAGCGATTAAT 899
Db 1014 TTAAGAGATACCTTTACGAAAGTAGTGAAATTCAGCAAGCTGAAAGAGTTCAAAATGATATC 955
Qy 900 AATTAATTTATTTCAATCTTAAGCCTGTTAAAGATATATCTGCAATTAGACTTGAATGCACT 959
Db 954 GCCAATCTGTTCCATTTTAAATCTTTAAGTGATATATCTGCAATTTGAATTTAAATGCACT 895
Qy 960 TTAGGTAAATTTCAAGAGATATTTATTTATGATGAGCATATAATAATAATTTATTTGAGACA 1019
Db 894 TTAGAAATTTTCAAGAAATTAATCTGATGATGATTAACGGTAGAAGTTAATTTCAACA 835
Qy 1020 TCGAATGATAACACAGATGAGAGTTGAACGAGTTTATGACACCGTTATTTTGACCGCTGA 1079
Db 834 TCAATGATAATACACTTGTCTTATGATAACAAAATTTGATTTCAAACATCTCTGAACGTATA 775
Qy 1080 ATAAAAAACCGTATAAAGCAATTTGAATTTAATTTAAAGAACCAATTTACAACGCA 1139
Db 774 CATATTCAAAGAACCCACGCTATTAATTTACTTAGTAATTAACGACCTATACGTTGCAAA 715
Qy 1140 GATTTCAAAGGATAGCTTGTAAATTCATTTCACTAGAAAATTTATGATAACATCTGTA 1199
Db 714 GATTTTCTGATACAGTATATAGTCCATTTCTTCAAAATTTATGATTAATCTCGTTAAA 655
Qy 1200 TCATTTGATATCATTTGCGCTGGCATTTGGAGTGATTTGCAACAATTTAATCTGCCAATC 1259
Db 654 TCACCTTATATCGTTGCACTTGTGTTGATTAATTTGCAACCATTTATCTGCGCGCTG 595
Qy 1260 AGTTATGATTTTCAACAATAATTTACTAAACCGCTTGTGAGTTTATCAAAATAAATGAT 1319
Db 594 AGTTATATCTTTCTTTCGCAAAATTTACTAAACCGATAGTTACAATGTCCAATTAATAAT 535
Qy 1320 GAGATTCGACGAGATGGTTTTCAAATAAAATTTGCAATTTAAATTAATAATTTAGAAATA 1379
Db 534 CAAATTAAGAGAGATGGTTTTCAAATAAATTTGAAATTAATTAATAATTTAGAAATA 475
Qy 1380 GATAATTTAGCAAAATACGTTTTAATGAGATGATGAGCCAAATTTGAAGAAATTTAATCAA 1439
Db 474 GATAATTTAATGATACCTTTTAAATGAATGATGATCAATAGAAGATCTTTAATCAG 415
Qy 1440 CAAAGACAAATTTGTTGAAGATCGGTCAATGAAATTTAGCAACCAATTTAATAATTTA 1499
Db 414 CAAAGCAATTTGTCGAGGATGCTTCAACGAAATTAAGAACGCCACTGCAGATTTAATCAA 355
Qy 1500 GGTCAATTTAAATTTGATTTACGAGTGGGAAAAAAGACCCAGCAGTATTAGAACAAATCG 1559
Db 354 GGTCAATTTAAATTTAATCCAACGTTGGGGAAAAAAGATCCAGCAGTTTTTGAAGAAATCT 295
Qy 1560 TTAATATTTCTATTGAAGAAATGAATCGTATCATATAAATTTAGTCCAAAGAAATTTACTGAA 1619
Db 294 TTGAATATTTCAATTTGAAGAGTGAATCGAATAACAAACTTTGCGAAGAACTACTTTTA 235
Qy 1620 TTGACTAAGAGAGATGTAATGACATTTCTTCTGAGCGCAGACCGTGCATTTAATGAT 1679
Db 234 CTTACCAAGATAGAGTCAATCAATTAATGTTTTGGAAATGTGAAAAATGTAGACGTAATAAGC 175

Search completed: October 5, 2004, 07:05:43
Job time : 8833.24 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 5, 2004, 00:12:59 ; Search time 903.032 Seconds
(without alignments)
10354.322 Million cell updates/sec

Title: US-09-006-627-1
Perfect score: 2201
Sequence: 1 TAAATTTAAAGCAACTATT.....AAACAAATAACAGTGGTT 2201

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 337363 seqs, 212409041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_29Jan04.*
1: Geneseq1980s.*
2: Geneseq1990s.*
3: Geneseq2000s.*
4: Geneseq2001as.*
5: Geneseq2001bs.*
6: Geneseq2002s.*
7: Geneseq2003as.*
8: Geneseq2003bs.*
9: Geneseq2003cs.*
10: Geneseq2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2201	100.0	2201	2	AAX90601
2	1372	62.3	1440	2	AAV75036
3	1356	61.6	1356	7	ACA48278
4	1349.8	61.3	1353	7	ACF74592
5	1085	49.3	3230	4	AAH54308
6	1041.6	47.3	3018	4	AAH54920
7	978.8	44.5	3099	4	AAH54422
8	736	33.4	736	2	AAT72328
9	736	33.4	736	2	AAX90602
10	732.8	33.3	1416	6	ABN90850
11	673.2	30.6	4858	2	AAV74690
12	657	29.9	657	7	ACF74591
13	471	21.4	855	4	AAH53016
14	471	21.4	855	4	AAH53698
15	401.8	18.3	669	6	ABN90912
16	399.8	18.2	110000	6	ABQ69245_14
17	399.8	18.2	110000	6	ABQ67195_3
18	369.8	16.8	110000	6	ABA03041_13
19	369.8	16.8	110000	6	ABA03041_14
20	302.6	13.7	501	4	AAH53697
21	302.6	13.7	501	4	AAH53015
22	246.6	11.2	1452	7	ACA36882
23	223.8	10.2	110000	6	ABA90521_16

c	24	212.8	9.7	110000	6	ABN71527_15	Continuation (16 o
	25	207.6	9.4	110000	6	ABN71527_20	Continuation (21 o
	26	206.2	9.4	1566	9	ADC93150	Adc93150 E. faeciu
c	27	197.4	9.0	1551	2	AAX13409	Aax13409 Enterococ
c	28	197.4	9.0	1551	6	ABN99204	ABN99204 Enterococ
c	29	194.8	8.9	1908	2	AAX12964	Aax12964 Enterococ
c	30	194.8	8.9	1908	6	ABN98759	ABN98759 Enterococ
	31	193	8.8	687	6	ABN68466	ABN68466 Streptoco
	32	187.2	8.5	687	9	ADC93255	Adc93255 E. faeciu
	33	184.2	8.4	110000	6	ABQ69245_13	Continuation (14 o
	34	175.2	8.0	110000	8	ADB12064_08	Continuation (9 of
	35	173.8	7.9	684	6	ABN67673	ABN67673 Streptoco
	36	161	7.3	705	8	ADB09375	ADB09375 Alloiooco
c	37	156.2	7.1	110000	6	ABN71527_06	Continuation (7 of
	38	155.2	7.1	110000	6	ABA03041_26	Continuation (27 o
c	39	154.8	7.0	110000	6	ABQ67195_5	Continuation (6 of
c	40	154.8	7.0	110000	6	ABQ69245_27	Continuation (28 o
	41	150.6	6.8	1581	8	ADB09377	ADB09377 Alloiooco
	42	150.6	6.8	1581	8	ADB09379	ADB09379 Alloiooco
c	43	149.8	6.8	741	6	ABN90875	ABN90875 Staphyloc
c	44	148.8	6.8	110000	6	ABN71527_09	Continuation (10 o
c	45	148.2	6.7	110000	6	ABA90521_17	Continuation (18 o

ALIGNMENTS

RESULT 1
AAX90601
ID AAX90601 standard; DNA; 2201 BP.

XX AC AAX90601;

XX DT 01-OCT-1999 (first entry)

XX DE Histidine kinase encoding polynucleotide sequence.

XX KW Histidine kinase; Two component signal transduction system; TCSTS;
virulence; bacterial pathogenesis; autophosphorylate; immunisation;
antibacterial agent; cognate response regulator; screening;
XX KW Helicobacter pylori; gastric ulcer; ds.
XX OS Staphylococcus aureus.

XX FH Key Location/Qualifiers
XX CDS 744..2099

XX FT /*tag= a
XX FT /product= "Histidine kinase protein"
XX FT /note= "Has sequence homology to Kina from Lactococcus
lactis cremoris"

XX FT WO9936508-A1.

XX PD 22-JUL-1999.

XX PF 12-JAN-1999; 99WO-US000610.

XX PR 13-JAN-1998; 98US-00006627.

XX XX (SMIK) SMITHKLINE BEECHAM CORP.

XX PA Wallis NG, Shilling LK, Mooney JL, Debouck C, Zhong Y;

XX PI Jaworski DD, Wang M, Throup JP;

XX XX WPI; 1999-444390/37.

XX DR P-PSDB; AAY28601.

XX XX Novel histidine kinase polynucleotides and polypeptides used to screen
for antibacterial compounds.

XX PS Claim 6; Page 5-6; 43pp; English.

XX CC The present sequence is a histidine kinase encoding novel polynucleotide

CC isolated from *Staphylococcus aureus*. The sequence was obtained from a
 CC library of chromosomal DNA clones of *Staphylococcus aureus* in E.coli.
 CC Histidine kinase is a component of the two component signal transduction
 CC system(TCSTS) which is involved in bacterial pathogenesis and virulence.
 CC It undergoes autophosphorylation at a histidine residue and phosphate
 CC group is transferred to the cognate response regulator. The nucleotide
 CC sequence can be used for diagnosis and staging of diseases, and as
 CC reagents for screening genetic mutations. The polynucleotide can also be
 CC used for therapeutic or prophylactic purposes and in particular for
 CC genetic immunisation. Polypeptides derived from this sequence, can be used
 CC to produce antibodies and to identify agonists and antagonists which are
 CC used to prevent, inhibit or treat diseases, particularly *Helicobacter*
 CC *pylori* infections, such as gastric ulcers, gastrointestinal carcinoma,
 CC and gastritis. The histidine kinase products can be used to screen new
 CC antibacterial drugs effective against resistant *S.aureus* strains
 XX
 SQ Sequence 2201 BP; 863 A; 290 C; 372 G; 676 T; 0 U; 0 Other;

Query Match 100.0%; Score 2201; DB 2; Length 2201;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAAATTTAAAGCAACATTTGTATAGAAAATACAAAATTTAAATATATTACCTTATTA 60
 Db 1 TAAATTTAAAGCAACATTTGTATAGAAAATACAAAATTTAAATATATTACCTTATTA 60

Qy 61 GAAAAAGTCGTATATAGAGGTGTACAAATGACGCAAAATTTTATAGTAGAAGATGAACAA 120
 Db 61 GAAAAAGTCGTATATAGAGGTGTACAAATGACGCAAAATTTTATAGTAGAAGATGAACAA 120

Qy 121 AACTTACCAAGATTTCTTGAATTTGGAATTCACATGAAATTTACAAATGTTGGACACAG 180
 Db 121 AACTTACCAAGATTTCTTGAATTTGGAATTCACATGAAATTTACAAATGTTGGACACAG 180

Qy 181 TATGATCGAACAAGCGTTTAGATAAAAGCGCTTAGCCATTTACTATGATTTTAAATCATATTA 240
 Db 181 TATGATCGAACAAGCGTTTAGATAAAAGCGCTTAGCCATTTACTATGATTTTAAATCATATTA 240

Qy 241 GATTTAATGTTGGCGTCAATTAATGGCTTAGAAATTTGTCGCAAAATTTAGACAAACAA 300
 Db 241 GATTTAATGTTGGCGTCAATTAATGGCTTAGAAATTTGTCGCAAAATTTAGACAAACAA 300

Qy 301 TCTACACCTATCATTAATAATTACAGCGAAAAGTGATACGATGACAAAGTTGCTGGGCTT 360
 Db 301 TCTACACCTATCATTAATAATTACAGCGAAAAGTGATACGATGACAAAGTTGCTGGGCTT 360

Qy 361 GATTACGGTGCAGACGATTTATAGTTAAGCGGTTTGATATTGAAGAACTTTTAGCAAGA 420
 Db 361 GATTACGGTGCAGACGATTTATAGTTAAGCGGTTTGATATTGAAGAACTTTTAGCAAGA 420

Qy 421 ATTGTCGAATTTTACGTCGTACGCCACAAAGGATATTATCGATGTCACCGGTATTACA 480
 Db 421 ATTGTCGAATTTTACGTCGTACGCCACAAAGGATATTATCGATGTCACCGGTATTACA 480

Qy 481 ATTGATAAGAACGCTTTTAAAGTGACGCTTAAATGGCGCAGAAATTTGAATTTAAACAAACA 540
 Db 481 ATTGATAAGAACGCTTTTAAAGTGACGCTTAAATGGCGCAGAAATTTGAATTTAAACAAACA 540

Qy 541 GAGTATGATTTTACTATATCTTTAGCTGAAAATTAACCAATGTTATGCAACGGGAACAA 600
 Db 541 GAGTATGATTTTACTATATCTTTAGCTGAAAATTAACCAATGTTATGCAACGGGAACAA 600

Qy 601 ATTTTAAATCATGTATGGGTTTAAATAGTAGAAGTAGAAAACAAATGTCGTAGATGTTTAT 660
 Db 601 ATTTTAAATCATGTATGGGTTTAAATAGTAGAAGTAGAAAACAAATGTCGTAGATGTTTAT 660

Qy 661 ATAAGATATTTACGAAACAAAGTTTAAACCAATGATGATGATGATGATGATGATGATGAT 720
 Db 661 ATAAGATATTTACGAAACAAAGTTTAAACCAATGATGATGATGATGATGATGATGATGAT 720

Qy 721 CGTGGCGTTGGGTATGTCATCGATGACAAAACGTAATTTGCGCAATTAACCTGGAATTATG 780
 Db 721 CGTGGCGTTGGGTATGTCATCGATGACAAAACGTAATTTGCGCAATTAACCTGGAATTATG 780

Qy 781 TTACCACGATGATTAGCTTTGTACGATATTTTGTGTTTGTGTTTAAATATATTATTTTCT 840
 Db 781 TTACCACGATGATTAGCTTTGTACGATATTTTGTGTTTGTGTTTAAATATATTATTTTCT 840

Qy 841 TGAAGATACACTGCATATAGTAGCTTTGATGATGCAGAACGAACTCAAGCGATATTA 900
 Db 841 TGAAGATACACTGCATATAGTAGCTTTGATGATGCAGAACGAACTCAAGCGATATTA 900

Qy 901 ATAATTTATTTCAATTTCTAAGCTGTTTAAAGATATATCTGCATTTAGACTTTGAATGATCTT 960
 Db 901 ATAATTTATTTCAATTTCTAAGCTGTTTAAAGATATATCTGCATTTAGACTTTGAATGATCTT 960

Qy 961 TAGGTAAATTTTCAAGAGATAATTTATTTATGATGATGATGATGATGATGATGATGATGAT 1020
 Db 961 TAGGTAAATTTTCAAGAGATAATTTATTTATGATGATGATGATGATGATGATGATGATGAT 1020

Qy 1021 CGAATGATTAACACAGTGAGAGTTTGAACACAGGTTTATGAACACCGCTTATTTTACCGGTA 1080
 Db 1021 CGAATGATTAACACAGTGAGAGTTTGAACACAGGTTTATGAACACCGCTTATTTGACCGGTA 1080

Qy 1081 TAAAAAAGCGTATAAAGCGATTTGAATTTTAAATTTTAAAGAACCAATTTACAAACGCAAG 1140
 Db 1081 TAAAAAAGCGTATAAAGCGATTTGAATTTTAAATTTTAAAGAACCAATTTACAAACGCAAG 1140

Qy 1141 ATTTCAAGGGTATAGCTTTGTTAAATTTCACTAGAAAATTTATGATTAACATCGTAAAAAT 1200
 Db 1141 ATTTCAAGGGTATAGCTTTGTTAAATTTCACTAGAAAATTTATGATTAACATCGTAAAAAT 1200

Qy 1201 CATTGATATCAATGCGCTGGCATTTGGAGTGATTCACCAATTTATAAAGTCCACAACTCA 1260
 Db 1201 CATTGATATCAATGCGCTGGCATTTGGAGTGATTCACCAATTTATAAAGTCCACAACTCA 1260

Qy 1261 GTTATGATTTTCAACACAAATTTACTAAACCGCTTCTCAGTTTATCAAAATAAAATGATTG 1320
 Db 1261 GTTATGATTTTCAACACAAATTTACTAAACCGCTTCTCAGTTTATCAAAATAAAATGATTG 1320

Qy 1321 AGATTCGACGAGATGGTTTTCAAAATAAAATTTCAAAATAAAATTTAGAAATGATG 1380
 Db 1321 AGATTCGACGAGATGGTTTTCAAAATAAAATTTCAAAATAAAATTTAGAAATGATG 1380

Qy 1381 ATAAATTTAGCAATACGTTTAAATGATGATGATGATGATGATGATGATGATGATGATGAT 1440
 Db 1381 ATAAATTTAGCAATACGTTTAAATGATGATGATGATGATGATGATGATGATGATGATGAT 1440

Qy 1441 AAAGCAATTTGTTGAAGATGCGTCACTGAATTTAGCAACCAATTTACAAATTTATTTCAAG 1500
 Db 1441 AAAGCAATTTGTTGAAGATGCGTCACTGAATTTAGCAACCAATTTACAAATTTATTTCAAG 1500

Qy 1501 GTCATTTAAATTTGATTCAGCGATGGGAAAAGAACCCAGCAGTATTAGAAAGATCGT 1560
 Db 1501 GTCATTTAAATTTGATTCAGCGATGGGAAAAGAACCCAGCAGTATTAGAAAGATCGT 1560

Qy 1561 TAAATTTCTATTGAAGAAATGAATTCGTATCATATAAAATTTAGTCGAAGAAATTTACTTGAAT 1620
 Db 1561 TAAATTTCTATTGAAGAAATGAATTCGTATCATATAAAATTTAGTCGAAGAAATTTACTTGAAT 1620

Qy 1621 TGACTAAAGAGATGTAATGATGATTTCTTCTGAAGCGGACGCGTGCATATTAAATGATG 1680
 Db 1621 TGACTAAAGAGATGTAATGATGATTTCTTCTGAAGCGGACGCGTGCATATTAAATGATG 1680

Qy 1681 AATTCGCTCGGAAATACACTCAATTAACCAATTTGATGATGATGATGATGATGATGATGATG 1740
 Db 1681 AATTCGCTCGGAAATACACTCAATTAACCAATTTGATGATGATGATGATGATGATGATGATG 1740

Qy 1741 ATCTGACATCTAAAAATCTAGAAAATTTAAATGAAACCTCATCAATTTGCAACAAATTTATTT 1800
 Db 1741 ATCTGACATCTAAAAATCTAGAAAATTTAAATGAAACCTCATCAATTTGCAACAAATTTATTT 1800

Qy 1801 TAACTTTTATGATTAATGCAATCAATATGATGATGATGATGATGATGATGATGATGATGATG 1860
 Db 1801 TAACTTTTATGATTAATGCAATCAATATGATGATGATGATGATGATGATGATGATGATGATG 1860

QY 1861 CAGGTTAAATAAGCAAAATAAATTGAAATTCAGATCATGGAATTTGGTATTCAG 1920
DB |||||||
DB 1861 CAGGTTAAATAAGCAAAATAAATTGAAATTCAGATCATGGAATTTGGTATTCAG 1920
QY 1921 AGGAAGATCAAGATTTCATTTTGGATCGCTTTTATCGAGTGGATAAATCTCGTTCAGAA 1980
DB |||||||
DB 1921 AGGAAGATCAAGATTTCATTTTGGATCGCTTTTATCGAGTGGATAAATCTCGTTCAGAA 1980
QY 1981 GTCAAGGCGGTAATGAGTCGGATTATCTATTGCTCAAAAAATCAATTCATTAACGGAG 2040
DB |||||||
DB 1981 GTCAAGGCGGTAATGAGTCGGATTATCTATTGCTCAAAAAATCAATTCATTAACGGAG 2040
QY 2041 GATCATTAAATAAATTAAGAGTGAATTAATAAGGAACAACGTTTAAATCATATTTAAT 2100
DB |||||||
DB 2041 GATCATTAAATAAATTAAGAGTGAATTAATAAGGAACAACGTTTAAATCATATTTAAT 2100
QY 2101 CATGCTCAGAGCTCAATCAAGTCATAGGATCAATTTTAAAGTACACATTAAGCTGTGA 2160
DB |||||||
DB 2101 CATGCTCAGAGCTCAATCAAGTCATAGGATCAATTTTAAAGTACACATTAAGCTGTGA 2160
QY 2161 CTAATGTATAAGAACAACTATATAAACAATAAACAAGTGGTT 2201
DB |||||||
DB 2161 CTAATGTATAAGAACAACTATATAAACAATAAACAAGTGGTT 2201

RESULT 2

AAV75036
ID: AAV75036 standard; DNA; 1440 BP.
XX
AC AAV75036;
XX
DT 16-MAR-1999 (first entry)
XX
DE Staphylococcus aureus contig SEQ ID #725.
XX
KW Computer readable medium; vaccine; S.aureus infection; immunodetection;
KW cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;
KW skin infection; surgical wound infection; scalded skin syndrome;
KW toxic shock syndrome; ds.
XX
OS Staphylococcus aureus.
XX
XX
Key Location/Qualifiers
FT 1021..1080
FT misc_feature. /*tag= a
FT /note= "these bases represent a line of missing text in
FT the sequence listing in the specification. They are
FT included to maintain the nucleotide numbering given in
FT the specification for this DNA sequence"
XX
XX
XX EP786519-A2.
XX
XX
XX 30-JUL-1997.
XX
XX
XX 07-JAN-1997; 97EP-00100117.
XX
XX
XX 05-JAN-1996; 96US-0009861P.
XX
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Kunsch CA, Choi GH, Baragh SC, Dillon PJ, Fannon MR, Rosen CA;
XX
XX WPI; 1997-374922/35.
XX
XX Polynucleotide(s) and proteins derived from Staphylococcus aureus -
XX stored on computer readable medium and used in the production of anti-
XX S.aureus vaccines.
XX
XX Claim 1; Page 1628-1629; 3271pp; English.
XX
XX This sequence represents one of 5191 Staphylococcus aureus DNA sequences
XX of the invention. The DNA sequences are recorded on a computer readable
XX medium, preferably selected from a floppy or hard disk, random access

CC memory (RAM), read-only memory (ROM) or CD-ROM. Homology searches using
CC the S.aureus DNA sequences allows putative functions to be assigned so
CC that protein-encoding or regulatory regions of commercial, therapeutic or
CC industrial importance can be obtained. Specifically, sequences which are
CC likely to encode antigens have been identified and these polypeptides can
CC be used in a vaccine composition against S.aureus infection. The
CC polypeptides can also be used in a kit for the immunodetection of
CC S.aureus in a sample. S.aureus is implicated in numerous human diseases,
CC including cellulitis, eyelid infections, food poisoning, osteomyelitis,
CC skin and surgical wound infections, scalded skin syndrome, toxic shock
CC syndrome, etc. Organisms transformed with the DNA sequences can be used
CC for recombinant production of the polypeptides. The new DNA sequences
CC (and their fragments) are useful as primers or probes for isolating
CC homologues of any of the S.aureus DNA sequences contained on the computer
CC readable medium
XX
SQ Sequence 1440 BP; 519 A; 179 C; 241 G; 437 T; 0 U; 64 Other;
Query Match 62.3%; Score 1372; DB 2; Length 1440;
Best Local Similarity 95.5%; Pred. No. 8.7e-242;
Matches 1372; Conservative 1; Mismatches 64; Indels 0; Gaps 0;
QY 29 AAATACAAAATTTAAATATATATATACCTTATTAGAAAGTCGTAATGAGGTGTACAAA 88
DB |||||||
DB 4 AAATNCANTATTTNAAATATATATATACCTTATTAGAAAAGTCGTAATGAGGTGTACAAA 63
QY 89 TGACGCAAAATTTAATAGTAGAAGATGAACAAAACCTTAGCAAGATTTCTTGAATTTGGAAC 148
DB |||||||
DB 64 TGACGCAAAATTTAATAGTAGAAGATGAACAAAACCTTAGCAAGATTTCTTGAATTTGGAAC 123
QY 149 TCACATCAAAAATTTCAATCTGGACACAGAGTATGATGGAACAAGCGTTTGTAGATAAG 208
DB |||||||
DB 124 TCACATCAAAAATTTCAATCTGGACACAGAGTATGATGGAACAAGCGTTTGTAGATAAG 183
QY 209 CGCTTAGCCATTACTATGATTTAATCATATATAGATTAACTGTCGCCGTCATTAATGCGT 268
DB |||||||
DB 184 CGCTTAGCCATTACTATGATTTAATCATATATAGATTAACTGTCGCCGTCATTAATGCGT 243
QY 269 TAGAAAATTTGTCGCAAAATTTAGACAAACAATCTACACCTATCATTAATAATTACAGCGA 328
DB |||||||
DB 244 TAGAAAATTTGTCGCAAAATTTAGACAAACAATCTACACCTATCATTAATAATTACAGCGA 303
QY 329 AAAGTGATACGTATGACAAAAGTTGCTGGCTTGATTAACGGTGCAGACGATTATATAGTTA 388
DB |||||||
DB 304 AAAGTGATACGTATGACAAAAGTTGCTGGCTTGATTAACGGTGCAGACGATTATATAGTTA 363
QY 389 AGCGTTTGATTTGAAGAATTTTACGAAGATTTTACGAAGATTTTACGTCGTGACGCCAC 448
DB |||||||
DB 364 AGCGTTTGATTTGAAGAATTTTACGAAGATTTTACGAAGATTTTACGTCGTGACGCCAC 423
QY 449 AAAAGGATATTATCGATGTCACACGGTATTACAATTTGATGAACACGCTTTTAAAGTGACGG 508
DB |||||||
DB 424 AAAAGGATATTATCGATGTCACACGGTATTACAATTTGATGAACACGCTTTTAAAGTGACGG 483
QY 509 TAAATGGCGCAGAAAATTTGAATTAACAAAACAGAGTATGATTTACTATATCTTCTAGCTG 568
DB |||||||
DB 484 TAAATGGCGCAGAAAATTTGAATTAACAAAACAGAGTATGATTTACTATATCTTCTAGCTG 543
QY 569 AAAATAAAACCATGTTATGCAACGGGAACAAAATTTTAAATCATGTATGGGGTTATAATA 628
DB |||||||
DB 544 AAAATAAAACCATGTTATGCAACGGGAACAAAATTTTAAATCATGTATGGGGTTATAATA 603
QY 629 GTGAAGTAGAACAACAAATGCTAGATGTTTATAGATATTTAAGATATTTACGAACAAAGTTTAAAC 688
DB |||||||
DB 604 GTGAAGTAGAACAACAAATGCTAGATGTTTATAGATATTTAAGATATTTACGAACAAAGTTTAAAC 663
QY 689 CATACGATCGTGACAAAATTTGAATTAACAAAACAGTTCGTCGGCTTGGGTATGTGATACGATAC 748
DB |||||||
DB 664 CATACGATCGTGACAAAATTTGAATTAACAAAACAGTTCGTCGGCTTGGGTATGTGATACGATAC 723
QY 749 AAAACGTAATTTGGCGCAATAACTGGATTTTGTACACAGATGATTAAGTTTGTACACGAT 808
DB |||||||
DB 724 AAAACGTAATTTGGCGCAATAACTGGATTTTGTACACAGATGATTAAGTTTGTACACGAT 783

QY 809 ATTTTGGTTTGGTTTAAATTAATTTTCTTCTGAAAGATACACTGCATTAATAGTGAGCT 868
Db 784 ATTTTGGTTTGGTTTAAATTAATTTTCTTCTGAAAGATACACTGCATTAATAGTGAGCT 843
QY 869 TGATGATGACAGAACCAAGCTCAAGCGATATTAATAATTTTCAATCTAAGCGCTGTAA 928
Db 844 TGATGATGACAGAACCAAGCTCAAGCGATATTAATAATTTTCAATCTAAGCGCTGTAA 903
QY 929 AGATATATCTGCATTAAGTCACTTTAGTGAATCTTTAGTGAATTTTCAAGAGATAATATTATTA 988
Db 904 AGATATATCTGCATTAAGTCACTTTAGTGAATCTTTAGTGAATTTTCAAGAGATAATATTATTA 963
QY 989 TGATGACATTAATAATAATTTTGGAGATCGAATGAATGAACACAGTGAGTTGAAC 1048
Db 964 TGATGACATTAATAATAATTTTGGAGATCGAATGAATGAACACAGTGAGTTGAANN 1023
QY 1049 AGGTTATGAACACCGTTATTTGACCGCTTAATAAAAAACGCTATAAGCGCATTTGAATA 1108
Db 1024 NNATA 1083
QY 1109 TTTAATTTAAGAACCAATTAACACGCAAGATTTCAAAAGGTATAGCTTGTAAATCA 1168
Db 1084 TTTAATTTAAGAACCAATTAACACGCAAGATTTCAAAAGGTATAGCTTGTAAATCA 1143
QY 1169 TTCACTAGAAAATTTATGATTAACATCGTAAATCATTTGATATCATTTGGCGCTGGCATTTGG 1228
Db 1144 TTCACTAGAAAATTTATGATTAACATCGTAAATCATTTGATATCATTTGGCGCTGGCATTTGG 1203
QY 1229 AGTGATTGCAACAAATTAATCTGCGCAATCAGTTATGTATTTTCAACACAAATTAATA 1288
Db 1204 AGTGATTGCAACAAATTAATCTGCGCAATCAGTTATGTATTTTCAACACAAATTAATA 1263
QY 1289 ACCGCTTGTGAGTTTCAATAAATAATGATGAGATTCGACGAGATCGTTTCAAAATAA 1348
Db 1264 ACCGCTTGTGAGTTTCAATAAATAATGATGAGATTCGACGAGATCGTTTCAAAATAA 1323
QY 1349 ATTGCAATTAATAAATAATGAGAAATAGATAATTTAGCAATACGTTTAATGAGAT 1408
Db 1324 ATTGCAATTAATAAATAATGAGAAATAGATAATTTAGCAATACGTTTAATGAGAT 1383
QY 1409 GATGAGCAAAATTTGAAGAATCATTTAATCAACAAAGCAATTTTGTGAAGATGGTC 1465
Db 1384 GATGAGCAAAATTTGAAGAATCATTTAATCAACAAAGCAATTTTGTGAAGATGGTC 1440

RESULT 3

ACA46278
ID ACA46278 standard; DNA; 1356 BP.

AC ACA46278;

XX 19-JUN-2003 (first entry)

XX Prokaryotic essential gene #27935.

XX Antisense; ds; prokaryotic essential gene; cell proliferation;
XX drug design; gene.

XX Staphylococcus aureus.

XX W0200277183-A2.

XX 03-OCT-2002.

XX 21-MAR-2002; 2002WO-US009107.

XX 21-MAR-2001; 2001US-00915242.

PR 06-SEP-2001; 2001US-00948993.

PR 25-OCT-2001; 2001US-0342923P.

PR 08-FEB-2002; 2002US-00072851.

PR 06-MAR-2002; 2002US-0362699P.

(ELIT-) ELITRA PHARM INC.

XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GU, Yamamoto R, Forsyth RA, Xu HH;

XX WPI; 2003-029926/02.
DR P-PSDB; ABU42408.

XX New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.

PS Claim 14; SEQ ID NO 34148; 1766pp; English.

XX The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC on a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. Note: The present sequence is one of the target
CC prokaryotic essential genes. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 1356 BP; 532 A; 178 C; 218 G; 428 T; 0 U; 0 Other;

XX Query Match 61.6%; Score 1356; DB 7; Length 1356;
XX Best Local Similarity 100.0%; Pred. No. 7.3e-239; Mismatches 0;
XX Matches 1356; Conservative 0; Indels 0; Gaps 0;

QY 744 ATGACAAAACGTAATTCGCGCAATACTGGATTATTGTTACCAGATGATTACGTTTGTGTC 803

Db 1 ATGACAAAACGTAATTCGCGCAATACTGGATTATTGTTACCAGATGATTACGTTTGTGTC 60

QY 804 ACGAATTTTGTGTTGTTAAATTAATTAATTTTCTTGAAGATACACTGCATAAATAGT 863

Db 61 ACGAATTTTGTGTTGTTAAATTAATTAATTTTCTTGAAGATACACTGCATAAATAGT 120

QY 864 GAGCTTGATGATGACAGAACGAGCTCAAGCGATTAATAATTTTCAATCTTAAGCGCT 923

Db 121 GAGCTTGATGATGACAGAACGAGCTCAAGCGATTAATAATTTTCAATCTTAAGCGCT 180

QY 924 GTTAAAGATATATCTGCATTAGACTTGAATGCATCTTTAGGTAAATTTTCAAGAGATAAT 983

Db 181 GTTAAAGATATATCTGCATTAGACTTGAATGCATCTTTAGGTAAATTTTCAAGAGATAAT 240

QY 984 ATTTATGATGAGCATATAATAATTTTGAAGATCGAATGATATACACAGTGAGGTT 1043

Db 241 ATTTATGATGAGCATATAATAATTTTGAAGATCGAATGATATACACAGTGAGGTT 300

QY 1044 GAACGAGTTATGAACACCGCTTATTTTGACCGCGTAATAAAAAACGCTATAAAGGCATT 1103

Db	301	GAAC	CGGTTATGAAC	CCGTTATTTT	TGACCGGGTAA	TAAAAA	ACGCTATA	AAAGCA	TT	360		
Qy	1104	GAAT	TATTTAA	TAAAGAA	CCAA	TATCA	AA	CGCA	GATTTCAA	AGGGTATAG	CTTGTTA	1163
Db	361	GAAT	TATTTAA	TAAAGAA	CCAA	TATCA	AA	CGCA	GATTTCAA	AGGGTATAG	CTTGTTA	420
Qy	1164	ATTCA	TTCACTAG	AAAAATTA	GTATAC	ATCAT	CGT	AAAA	TATGTATAT	CA	TTCGCTGGCA	1223
Db	421	ATTCA	TTCACTAG	AAAAATTA	GTATAC	ATCAT	CGT	AAAA	TATGTATAT	CA	TTCGCTGGCA	480
Qy	1224	TTTGG	AGTGAT	TGCA	CAAA	TATTA	AA	CTGCC	ACAA	TATCA	GTATTTTCA	1283
Db	481	TTTGG	AGTGAT	TGCA	CAAA	TATTA	AA	CTGCC	ACAA	TATCA	GTATTTTCA	540
Qy	1284	ACTAAA	CCGCTTGT	CAGTTTAT	CAAA	TAAAA	TGAT	TGAGAT	TCGA	CGAGAT	TGTTTCAA	1343
Db	541	ACTAAA	CCGCTTGT	CAGTTTAT	CAAA	TAAAA	TGAT	TGAGAT	TCGA	CGAGAT	TGTTTCAA	600
Qy	1344	AATAA	ATTGCA	TTTAA	TATAC	AAATAT	TGA	AGAA	TAGAT	TAATTTAG	CAAA	1403
Db	601	AATAA	ATTGCA	TTTAA	TATAC	AAATAT	TGA	AGAA	TAGAT	TAATTTAG	CAAA	660
Qy	1404	GAGAT	GATGAG	CCAA	TTTGA	AGAA	TCA	TTAA	TCA	CAAA	AGCAATTTG	1463
Db	661	GAGAT	GATGAG	CCAA	TTTGA	AGAA	TCA	TTAA	TCA	CAAA	AGCAATTTG	720
Qy	1464	TCACAT	GAA	TTTAC	GAA	CCCA	TATTA	CAAA	TATTA	CAAGGTCAT	TTAA	1523
Db	721	TCACAT	GAA	TTTAC	GAA	CCCA	TATTA	CAAA	TATTA	CAAGGTCAT	TTAA	780
Qy	1524	TGGG	AAAAA	AGAC	CCG	CAG	AGATTA	TGA	AGAA	TATTC	TGTAAG	1583
Db	781	TGGG	AAAAA	AGAC	CCG	CAG	AGATTA	TGA	AGAA	TATTC	TGTAAG	840
Qy	1584	AATCG	TATCA	TAAATTTAG	TCG	AGAA	TACT	TGA	TTG	ACTA	AA	1643
Db	841	AATCG	TATCA	TAAATTTAG	TCG	AGAA	TACT	TGA	TTG	ACTA	AA	900
Qy	1644	ATTT	CTTCTG	AA	CGCAG	ACCG	TATTA	TGAT	GAAA	TT	CGCTCG	1703
Db	901	ATTT	CTTCTG	AA	CGCAG	ACCG	TATTA	TGAT	GAAA	TT	CGCTCG	960
Qy	1704	TTAAA	CAAT	TGAT	CTCTG	ATTA	CA	ATTTG	TATCG	GATCT	AAAA	1763
Db	961	TTAAA	CAAT	TGAT	CTCTG	ATTA	CA	ATTTG	TATCG	GATCT	AAAA	1020
Qy	1764	ATTA	AAATG	AA	CC	TCA	TAA	TTG	CGAC	CAAT	TTTTT	1823
Db	1021	ATTA	AAATG	AA	CC	TCA	TAA	TTG	CGAC	CAAT	TTTTT	1080
Qy	1824	AAAT	TATGAT	GTGA	AGAA	TAA	GAAA	TTAAA	AGTTTAA	GACA	AGGTTAAAA	1883
Db	1081	AAAT	TATGAT	GTGA	AGAA	TAA	GAAA	TTAAA	AGGTTAA	GACA	AGGTTAAAA	1140
Qy	1884	ATA	TTGAA	ATTA	CAGAT	TCAT	TGGA	TTTCC	AGAG	AAGAT	CAAG	1943
Db	1141	ATA	TTGAA	ATTA	CAGAT	TCAT	TGGA	TTTCC	AGAG	AAGAT	CAAG	1200
Qy	1944	GAT	CGCTTTTAT	CGAG	TGGA	TAA	TCTCG	TTCA	AGAG	TCA	AGCGG	2003
Db	1201	GAT	CGCTTTTAT	CGAG	TGGA	TAA	TCTCG	TTCA	AGAG	TCA	AGCGG	1260
Qy	2004	TTAT	CTATT	TGCT	CA	AAAA	TATTA	CGG	AGGAT	TCGAT	TAAAA	2063
Db	1261	TTAT	CTATT	TGCT	CA	AAAA	TATTA	CGG	AGGAT	TCGAT	TAAAA	1320
Qy	2064	ATTA	ATA	AGGA	CA	ACG	TTTAAA	TATTA	TTTAA			2099
Db	1321	ATTA	ATA	AGGA	CA	ACG	TTTAAA	TATTA	TTTAA			1356

RESULT 4

ACF74592	ACF74592 standard; DNA; 1353 BP.
XX	AC
XX	AC
XX	ACF74592;
DT	20-NOV-2003 (first entry)
XX	
XX	Staphylococcus aureus DNA #2272.
DE	
XX	Antibacterial; vaccine; gene therapy; infection; sepsis; diagnosis;
KW	enzymatic assay; antibiotic target; gene; ds.
XX	
OS	Staphylococcus aureus.
XX	
PN	WO200294868-A2.
XX	
PD	28-NOV-2002.
XX	
PF	27-MAR-2002; 2002WO-IB002637.
XX	
PR	27-MAR-2001; 2001GB-00007661.
XX	
PA	(CHIR-) CHIRON SPA.
PI	Masignani V, Mora M, Scarselli M;
XX	
DR	WPI; 2003-120786/11.
DR	P-PSDB; ABM73032.
XX	
PT	New Staphylococcus aureus protein, useful as a vaccine for treating or
PT	preventing Staphylococcal infection, specifically an infection caused by
PT	S. aureus, e.g. sepsis.
XX	
PS	Claim 6; SEQ ID NO 4543; 49pp; English.
XX	
CC	The invention relates to novel genes and encoded proteins from
CC	Staphylococcus aureus. A composition comprising the S. aureus protein, a
CC	nucleic acid encoding the protein, or an antibody to the protein, is
CC	useful as a pharmaceutical, particularly as a vaccine for treating or
CC	preventing infection due to Staphylococcus bacteria, specifically an
CC	infection caused by S. aureus. The composition is particularly useful for
CC	treating or preventing sepsis in a patient. The composition can also be
CC	used for diagnostics. The protein is also used in an assay for enzymatic
CC	studies and as a target for antibiotics. This sequence represents one of
CC	the novel S. aureus genes of the invention
XX	
SQ	Sequence 1353 BP; 531 A; 179 C; 217 G; 426 T; 0 U; 0 Other;
Query Match 61.3%; Score 1349.8; DB 7; Length 1353;	
Best Local Similarity 99.9%; Pred. NO. 9.9e-238;	
Matches 1351; Conservative 0; Mismatches 2; Indels 0; Gaps 0;	
Qy	744 ATGACAAAACGTAAATTGGCAATACTGGATTATTGTTACCAAGATGATTACGTTTGTC 803
Db	1 ATGACAAAACGTAAATTGGCAATACTGGATTATTGTTACCAAGATGATTACGTTTGTC 60
Qy	804 ACCATATTTTGTGTTTGTATTAATTATTTTCTTGAAAGATACACTGCATAATAGT 863
Db	61 ACCATATTTTGTGTTTGTATTAATTATTTTCTTGAAAGATACACTGCATAATAGT 120
Qy	864 GAGCTTGTATGATGCAGAACGAAGCTCAAGCGATTAATTAATTTATTTCAATCTAAGCCT 923
Db	121 GAGCTTGTATGATGCAGAACGAAGCTCAAGCGATTAATTAATTTATTTCAATCTAAGCCT 180
Qy	924 GTTAAAGATATATCTGCATTAGACTTGAATGCATCTTTTAGGTAATTTTCAAGAGATAATT 983
Db	181 GTTAAAGATATATCTGCATTAGACTTGAATGCATCTTTTAGGTAATTTTCAAGAGATAATT 240
Qy	984 ATTTATGTAGGCATATAATAAATTTATTTTGACATCGAATCGAATCAACAGTGAGAGTT 1043
Db	241 ATTTATGTAGGCATATAATAAATTTATTTTGACATCGAATCGAATCAACAGTGAGAGTT 300
Qy	1044 GAACACGGTTATGAACACACGGTTATTTTGACCGCGTAAATAAAAAACGGTATAAAGGCATT 1103

```
Db 301 GAACAGGTTATGAACACCGTTATTTTACCGGTAAATAAAAAACGCTATAAGGCATT 360
Qy 1104 GAATATTTAATATTAAAGAACCAATTAACAACGCAAGATTTCAAGGGTATAGCTTGTTA 1163
Db 361 GAATATTTAATATTAAAGAACCAATTAACAACGCAAGATTTCAAGGGTATAGCTTGTTA 420
Qy 1164 ATTCATTCACCTAGAAAAATTATGATPAACATCGTAAAAATCATTTGTATATCATTTGGCTGGCA 1223
Db 421 ATTCATTCACCTAGAAAAATTATGATPAACATCGTAAAAATCATTTGTATATCATTTGGCTGGCA 480
Qy 1224 TTTGGAGTGTGCAACAATTTATAACTGCCACATCAGTTATGTATTTTCAACACAAAT 1283
Db 481 TTTGGAGTGTGCAACAATTTATAACTGCCACATCAGTTATGTATTTTCAACACAAAT 540
Qy 1284 ACTAAACCGCTGTGCAGTTTATCAATATAAATGATGAGATTCGACGAGATGGTTTCAA 1343
Db 541 ACTAAACCGCTGTGCAGTTTATCAATATAAATGATGAGATTCGACGAGATGGTTTCAA 600
Qy 1344 AATAAATTTGCAATTAATAACAAATTTATGAAGAAATAGATAATTTAGCAAAATAGCTTTAAT 1403
Db 601 AATAAATTTGCAATTAATAACAAATTTATGAAGAAATAGATAATTTAGCAAAATAGCTTTAAT 660
Qy 1404 GAGATGATGAGCCAAATTTGAAGAATCAATTTAATCAACAAGACAAATTTGTTGAAGTGG 1463
Db 661 GAGATGATGAGCCAAATTTGAAGAATCAATTTAATCAACAAGACAAATTTGTTGAAGTGG 720
Qy 1464 TCACATGAATTTAGCAACACCATTAACAATTTATCAAGTCATTTAATTTGATTCACGGA 1523
Db 721 TCACATGAATTTAGCAACACCATTAACAATTTATCAAGTCATTTAATTTGATTCACGGA 780
Qy 1524 TGGGAAAAAAGACCCAGCAGTATTAGAAATTCGTTAAATATTTCTATTGAAGAAATG 1583
Db 781 TGGGAAAAAAGACCCAGCAGTATTAGAAATTCGTTAAATATTTCTATTGAAGAAATG 840
Qy 1584 AATCGTATCAATAAATTTAGTCGAAGAATTAATCTGAATTTGACTAAAGGAGATGTAATGAC 1643
Db 841 AATCGTATCAATAAATTTAGTCGAAGAATTAATCTGAATTTGACTAAAGGAGATGTAATGAC 900
Qy 1644 ATTTCTTTCTAGCGCAGACCGTGCATATTAATGATGAATTCGCTCGCAATACACTCA 1703
Db 901 ATTTCTTTCTAGCGCAGACCGTGCATATTAATGATGAATTCGCTCGCAATACACTCA 960
Qy 1704 TTAACAATTTGCAATCTGATTTATCAATTTGATACGATCTGACATCTAAAAATCTAGAA 1763
Db 961 TTAACAATTTGCAATCTGATTTATCAATTTGATACGATCTGACATCTAAAAATCTAGAA 1020
Qy 1764 ATTAATAATGAACCTCATCAATTCGAACAATTTATTTTAAATCTTTTATGATAATGCAATC 1823
Db 1021 ATTAATAATGAACCTCATCAATTCGAACAATTTATTTTAAATCTTTTATGATAATGCAATC 1080
Qy 1824 AATATGATGTGAAGATAGAAAATTTAAAGTTAAGACAGGTTAAAAAATAGCAAAAA 1883
Db 1081 AATATGATGTGAAGATAGAAAATTTAAAGTTAAGACAGGTTAAAAAATAGCAAAAA 1140
Qy 1884 ATAAATCAAAATTACAGATCATGGAATTTGTTATTCAGAGGAAGATCAAGATTTTCATTTT 1943
Db 1141 ATAAATCAAAATTACAGATCATGGAATTTGTTATTCAGAGGAAGATCAAGATTTTCATTTT 1200
Qy 1944 GATCGCTTTTATCGAGTGGATAAATCTCGTTTCAAGAAAGTCAAGCGCGTAAATGGAATCGGA 2003
Db 1201 GATCGCTTTTATCGAGTGGATAAATCTCGTTTCAAGAAAGTCAAGCGCGTAAATGGAATCGGA 1260
Qy 2004 TTATCTATTCGCTCAAAAATCATTAATTAACCGGAGATCGATTAATAATTAAGTGAA 2063
Db 1261 TTATCTATTCGCTCAAAAATCATTAATTAACCGGAGATCGATTAATAATTAAGTGAA 1320
Qy 2064 ATTAATAAGGAACAACGTTTAAAAATCATATTT 2096
Db 1321 ATTAACAAGGAACAACGTTTAAAAATCATATTT 1353
```

```
AAH54308
ID AAH54308 standard; DNA; 3230 BP.
XX
AC AAH54308;
XX
DT 03-SEP-2001 (first entry)
XX
DE S. epidermidis genomic polynucleotide sequence SEQ ID NO:3672.
XX
KW Staphylococcus epidermidis SRI strain; infection; diagnosis; vaccination;
endocarditis; de.
XX
OS Staphylococcus epidermidis.
XX
PN WO200134809-A2.
XX
PD 17-MAY-2001.
XX
PF 09-NOV-2000; 2000WO-US030782.
XX
PR 09-NOV-1999; 99US-0164258P.
XX
PA (GLAX ) GLAXO GROUP LTD.
XX
PI Kimmerly WJ;
XX
PS WPI; 2001-316495/33.
XX
PT Nucleic acids encoding polypeptides from Staphylococcus epidermidis,
useful for vaccinating against infections, e.g. endocarditis.
XX
PS Claim 8; Page 1281-1282; 2188pp; English.
XX
CC AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides
(II), given in AA81454 to AG83120, from Staphylococcus epidermidis. (I)
and (II) can have antibacterial activity and therefore can be used in
vaccination. The nucleic acids (I) may be used to produce the S.
epidermidis polypeptides (II) via the production of vectors containing
them which are used to produce hosts cells which express the
polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
used to vaccinate subjects and to raise antibodies against the bacteria.
The polypeptides may also be used to assay for other inhibitors of their
activity and therefore identify compounds that may be used for the
treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to
AAH55090 represent specifically claimed S. epidermidis genomic DNA
polynucleotide sequences from the present invention. AAH55091 to AAH55098
represent oligonucleotide sequences and primers which are used in the
exemplification of the present invention. N.B. The present invention
specifically claims all the polynucleotide sequences given in the
sequence listing of the present specification, however the sequence
listing only goes up to SEQ ID NO:4454 so even though sequences are given
in the disclosure for SEQ ID NO:4465 to 4472, no sequences are present
for SEQ ID NO:4455 to 4464
XX
SQ Sequence 3230 BP; 1096 A; 524 C; 501 G; 1109 T; 0 U; 0 Other;
```

```
Query Match 49.3%; Score 1085; DB 4; Length 3230;
Best Local Similarity 71.7%; Pred. No. 3.1e-189;
Matches 1437; Conservative 0; Mismatches 565; Indels 3; Gaps 1;

Qy 95 AAATTTTAAATAGTAGAAGATGAACAAAACCTTAGCAAGATTTCTTGAATGGAACACTCACAC 154
Db 3 ATATTTTAAATTTGTAAGATGACGCAAAATCTTCTAGATTTATAGAGCTTGAGTTAACTC 62
Qy 155 ATGAAAATTAACAATGTGGACACAGAGTATGATGGACAACGCGTTTAGATAAAGCGCTTA 214
Db 63 ATGAAAATTAATACTGTTGATATAGAAATGATGAAAAGGTGGGTTTGGATAAAGCAATTAT 122
Qy 215 GCCATTACTATGATTTAATCATATTAGATTTATGTTTGGCTCAATTAATGGCTTGAGAA 274
Db 123 CTAAAGCCCTGATGATTTATATATTTTAGACTTAATGCTTCCAAACATAAATGGTCTGAAA 182
Qy 275 TTTGTGCGAAAATTAGACAACAACTATACCTATCATTTATATTTACAGCGAAAAGTG 334
```

Db 183 TTTGTAGACAAATTCGTCAAAAAACAACCTACTCAATTTATCATCTTACTGCAAAAAGCG 242
Qy 335 ATACGCTATGACAAAGTTGCTGGCTTGAATTACGGTGCAGACATTATATAGTTAAGCCGT 394
Db 243 AGACATATGATAAAGTAGCTGGTGGCTATGGGGCAGATGACTACATTGTAAACCCCT 302
Qy 395 TTGATATTGAAGAACTTTTGTAGCAAGAAATTCGTGCAATTTTACGTGTCAGCCACAAAAGG 454
Db 303 TTGATATAGAAGAAATTCGTGCAAGAAATTAAGAGCGGTATTCGCGCAGACAGCCAGATAAG 362
Qy 455 ATATTATCGATGTCACCGTATTACAAATTGATAAGAACGCTTTTAAAGTGACGTTAAATG 514
Db 363 ATGTTTTAGATATCAATGCTATTCATTGATAAAGATGCCCTTTAAAGTTACTGTTAATG 422
Qy 515 GCGCAGAAATTCGAATTAACAAAAACAGAGTATGATTTTACTATATCTTCTAGCTGAAAAATA 574
Db 423 GCCATCAATTAGAAATTAACAAAAACAGATACGATTTATATGTTTGTAGCTGAAATC 482
Qy 575 AAAACATGTTATGCAACGGGAACAAATTTTAAATCATGTATGGGGTTATATAGTGAAG 634
Db 483 GTAACCAAGTATGACAGCGTGAACAAATTCGATCAGTATGGGGTTATATAGTGAAG 542
Qy 635 TAGAACCAATTCGTAGATGTTTATATAGATATTTACGAAACAAAGTTAAACCATAGC 694
Db 543 TAGAAACGAATTCGTGATGTTTACATTCGTTATTTACGTAATAAATCTCAAAACCTTTTA 602
Qy 695 ATCTGTGACAAAAATGATGAACAGTTTCGTGGGTTGCGGTATGTCGATACGATCACAACAG 754
Db 603 ATAAGAAAAATCCATAGAAACAGTACGTGGCGTAGGGTATGTAATCGATGTTAAGCG 662
Qy 755 ---TAAATTCGCAATAACTGATTAATTTTACCAAGATGATACGTTTGTCAAGATAT 811
Db 663 CCAAAAAATTAATATGAATGATGCTTATTACGAGCTCATTACTTCAAGCAATAT 722
Qy 812 TTTGTTTGTGTTAATTTATTTTCTTTGAAAGATACATCGCATATATAGTACCTTGA 871
Db 723 ACTTTTCTGCTTAATTTATTTTCTTTTAAAGATACCTTTTACGAAGTAGTGAATTTGA 782
Qy 872 TGATCAGAACGAAGCTCAAGCGATTAATAATTTATTTTCAATCTAAGCCCTGTTAAAGA 931
Db 783 CGAAGCTGAAAGAGTTCAATGATATCGCCAACTCTGTTCCAACTCTTAATCTTTAAGTGA 842
Qy 932 TATATCTGCATTGACCTGAAATGATCTTTAGGTAAATTTTCAAGAGATAATTTATTTATGA 991
Db 843 TATATCTGCATTGGATTTAAATGATCTCTTAGAAAATTTTCAAGAAATCTGATCTATGA 902
Qy 992 TGAGCATATAATAATTTATTTGAGACATCGAATGATTAACACAGTGAAGTTGAACCCAGG 1051
Db 903 TGATAACCGTAGAAAAGTTAAATCAAAACATCAAAATGATAACACTTGCTTTATGATAACAA 962
Qy 1052 TTATCAACACCGTTATTTTGCCGGTAAATAAAAAACGCTATAAAGCATTTGATATTT 1111
Db 963 AATGATTTCAACATCTCGAAGCTATACATATTCAAAGAGCCACGGTATTAATTTACTT 1022
Qy 1112 AATTATTAAGAACCAATTTACAAACGAAGATTTCAAGGGTATAGCTTGTGTTAATTCATTC 1171
Db 1023 AGTAATACTGAACCTATACGTTCCGAAGATTTTCTGGATACAGTATATAGTCCATTC 1082
Qy 1172 ACTAGAAATTTATGATATACATCGTAAATCATTTGATATATATGTCGTCGCAATTTGGAGT 1231
Db 1083 TCTTCAAAATTTATGATAATCTCGTTAAATCACTTTATATCGTTGACATCTGCTTTGGATT 1142
Qy 1232 GATTCGAACAAATTTAATCGCCACAACTCAGTTATGTTTCAACACAAATTTACTTAACC 1291
Db 1143 AATTGCAACCATTTACTTCCCGGCGGTGATTTATCTTTTCTCGCAATTTACTTAACC 1202
Qy 1292 GCTTCTCAGTTTATCAATTAATGATTTGAGATTCGACGAGATGTTTCAAAATAAATTT 1351
Db 1203 GATAGTTACATGTTCAATTAATGAATCAATTAAGAGAGATGTTTCAAAATAAATTT 1262
Qy 1352 GCAATTAATAAATTTATGAAGAAATAGATAATTTAGCAATATAGTTTAAATGAGATCAT 1411

Db 1263 TGAATTAATACTAAATTTATGAAGAAACAGATAAATTTAATTTGATACTTTTAAATGAATGAT 1322
Qy 1412 GAGCCAAATTTGAAGAATCATTTAATCAACAAGACAAATTTGTTGAAGATGCGTCAATGA 1471
Db 1323 GTATCAATAGAGAAATCTTTTAAATCAGCAACGTCATTTGTCGAGGATGCTTTCACAGA 1382
Qy 1472 ATTAGCAACACATTTAATAATTTTCAAGGTCATTTTAAATTTTATGATTCAGCGATGGGAAA 1531
Db 1383 ATTAAGAACGCCACTGTCAGATTTATCAAGGTCATCTAAATTTAATCCAACTGGGGAAA 1442
Qy 1532 AAAAGACCCACAGATTTAGAGAAATCGTTAAATTTTCTATTGGAAGAAATGAATCGTAT 1591
Db 1443 AAAAGATCCAGAGTTTGGAGAAATCTTTGAAATTTTCAATTTGAAGAGTGAATCCGAAT 1502
Qy 1592 CATAAATTTAGTCCGAAGAAATTTACTTTGAATTCAGCTAAAGGAGATGTAAATGACATTTCTTC 1651
Db 1503 AACAAATCTGTCGAAGAACTACTTTTACTTACCAGAGATAGAGTCAATCATATGTTTT 1562
Qy 1652 TGAAGCGCAGACCGTGCATATTAATGATGAAATTCGCTCGCGAATACACTCATTTAAACA 1711
Db 1563 GGAATGTGAAGATGTAGACGTAATAGCGAGATTCATCACTGTAAGTCACTGCAACA 1622
Qy 1712 ATTCATCTCTGATTTCAATTTGATAGGATCTGACATCTAAATCTAGAAATTTAAAT 1771
Db 1623 CCTACATCCAGATTTACTTTTGAACACACATCTTGTACTAAGCTATCCAAATTTAAAT 1682
Qy 1772 GAAACCTCATCAATTCGAACAAATTTATTTTAAATCTTTATGATTAATGCAATCAAAATATGA 1831
Db 1683 TAACCGTATCAGTTTGAACAACTTACTCATATTTATGATTAATGCAATGAATATGA 1742
Qy 1832 TGTGAAGAAATTAAGAAATTTAAAGTTTAAAGCAAGGTTAAAGAAATTAAGCAAAATTAATGA 1891
Db 1743 CACTGAACATTAAGCACATTTAAATTTGTTACTCACTAAATAAATAAATGATTTATGATTGA 1802
Qy 1892 AATTACAGATCATGGAATTTGTTATTCAGAGAGAGATCAAGATTTCAATTTTGTATCGCTT 1951
Db 1803 TATTACTGATCATGGTATGGGTATACCAAAAGCTGACTTAGAATTTATCTTTGATAGATT 1862
Qy 1952 TTATCGATGATATAAATCTCGTTCAAGAAAGTCAAGCGGCTAATGAGCTCGGATTTATCTAT 2011
Db 1863 TTATCGTGTAGTAATAATCACGTCGTCGTAGTCAAGAGGCAATGGATTAGGACTATCAAT 1922
Qy 2012 TGCTCAAAAAATTCATTAATTTAAACGGAGGATCGATTTAAATTTAAAGTGAATTAATTAATA 2071
Db 1923 AGCAGAAAAATTTGTCAACTTAACGGTGGTATGATTCAGGTAGAAAGTGAAGTGAACATA 1982
Qy 2072 AGGAACAACGTTTAAATCATATTT 2096
Db 1983 GTACACGACTTTTCAAAATCAGTTTT 2007

RESULT 6

AAH54920/c
ID AAH54920 standard; DNA; 3018 BP.

XX AAH54920;

XX AC

XX 03-SEP-2001 (first entry)

XX S. epidermidis genomic polynucleotide sequence SEQ ID NO:4284.

XX Staphylococcus epidermidis SRI strain; infection; diagnosis; vaccination;

XX endocarditis; ds.

XX Staphylococcus epidermidis.

XX OS

XX WO200134809-A2.

XX PN

XX 17-MAY-2001.

XX PD

XX 09-NOV-2000; 2000WO-US030782.

XX PF

XX 09-NOV-1999; 99US-0164258P.

XX PR

XX PA (GLAX) GLAXO GROUP LTD.
 XX PI Kimmerly WJ;
 XX XX WPI; 2001-316495/33.
 XX PT Nucleic acids encoding polypeptides from *Staphylococcus epidermidis*,
 XX XX useful for vaccinating against infections, e.g. endocarditis.
 XX PS Claim 8; Page 2015-2016; 2189pp; English.
 CC AAH52304 to AAH53970 represent nucleic acids (i) encoding polypeptides
 CC (ii), given in AAG81454 to AAG83120, from *Staphylococcus epidermidis*. (i)
 CC and (ii) can have antibacterial activity and therefore can be used in
 CC vaccination. The nucleic acids (i) may be used to produce the S.
 CC epidermidis polypeptides (ii) via the production of vectors containing
 CC them which are used to produce hosts cells which express the
 CC polypeptides. The polypeptides (ii) (and/or nucleic acids) may then be
 CC used to vaccinate subjects and to raise antibodies against the bacteria.
 CC The polypeptides may also be used to assay for other inhibitors of their
 CC activity and therefore identify compounds that may be used for the
 CC treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to
 CC AAH55090 represent specifically claimed S. epidermidis genomic DNA
 CC polynucleotide sequences from the present invention. AAH55091 to AAH55098
 CC represent oligonucleotide sequences and primers which are used in the
 CC exemplification of the present invention. N.B. The present invention
 CC specifically claims all the polynucleotide sequences given in the
 CC sequence listing of the present specification, however the sequence
 CC listing only goes up to SEQ ID NO:4454 so even though sequences are given
 CC in the disclosure for SEQ ID NO:4465 to 4472, no sequences are present
 CC for SEQ ID NO:4455 to 4464
 XX XX
 SQ Sequence 3018 BP; 929 A; 508 C; 459 G; 1122 T; 0 U; 0 Other;
 Query Match 47.3%; Score 1041.6; DB 4; Length 3018;
 Best Local Similarity 71.5%; Pred. No. 2.6e-181;
 Matches 1384; Conservative 0; Mismatches 549; Indels 3; Gaps 1;
 QY 164 ACATGTTGGACACAGTATGTCGACAGACGGTTAGATTAAGCGCTTAGCCATTACT 223
 DB 3018 ATACTGTTGATATAGAAATGATGGAAGGTGGTGGTAAAGCAATATCTAAGCCCT 2959
 QY 224 ATGATTTAATCAVATTAGATTATGTTGGCGTCAATTAATGGCTTGAAGATTGTGCGCA 283
 DB 2958 ATGATTTATATTTTAGACTTAATGCTTCCAACATAAATGGTCTAGAAATTTGTAGAC 2899
 QY 284 AATTAGAACAAACAACTACACTATCAATTAATTAATTAAGCGAAAGTGATCGTATG 343
 DB 2898 AATTCGTCAAAAACAACTACTCCAATTTATCATCTACTGCAAAAAGCGAGACATATG 2839
 QY 344 ACAAAGTTGCTGGCTTGATTAACGTCAGACGATTTATATAGTTAAGCCGTTTGATATG 403
 DB 2838 ATAAAGTAGCTGGGTTGCGACTATGGGCGAGATGACTACATTTGTAACCCCTTTGATATAG 2779
 QY 404 AAGAACTTTTAGCAAGAAATTCGTCAATTTTACGTCGTGAGCCACAAAGGATATTATCG 463
 DB 2778 AAGAATTGCTCGCAAGATAAGAGCGGTATTCGACAGACAGCAGATAAGATGTTTAG 2719
 QY 464 ATGTCAACGGGTATTAACAATTGATAAGAACGCTTTTAAAGTGACGGTAAATGGCGCAGAAA 523
 DB 2718 ATATCAATGGTATTATCATTTGATAAAGATGCGTTTAAAGTTACTGTTAAATGGCCATCAAT 2659
 QY 524 TTGAATTAACAAACACAGGTATGATTTACTATATCTTCTGCTGAAATAAACCATG 583
 DB 2658 TAGAATTAACAAACACAGATACGATTTATATATTTTGTAGCTGAAATTCGTAACCCG 2599
 QY 584 TTATGCAACGGGAACAAATTTTAAATCATGCTATGGGGTTAATATAGTGAAGTAGAACAACAA 643
 DB 2598 TCAAGCAGCGTGAACAAATTTCTCGATCACTGATGGGGGTATATATAGTGAAGTAGAACAACGA 2539
 QY 644 ATGTCGTAGATGTTTATATAGATATTTTACGAAACAAAGTTTAAACCAATACGATCGTGACA 703
 DB 1578 TGTCGAAGAACTACTCTTTTACTTACCAGAGATAGAGTCAATCATTAATGTTTGGAAATGTGA 1519
 QY 1661 GACCGTGCATATTAAATGATGAAATTCGTCGCGAAATACACTCATTAAAAACAATTTGCATCC 1720
 DB 1518 AATGTAGACGTAAATAGCGAGATTCAATCACTGCTGAGTCACTGCAACACCTACATCC 1459
 QY 1721 TGATTAATCAATTTGATACGGATCTGACATCTAAAAATCTAGAAATCTAGAAATTAATAAGAAACCTCA 1780
 DB 1458 AGATTTACTTTTGAACACATCTTCTGCTACTAAGCCTATCTCAATTAATAAATAAACCCTCA 1399
 DB 2538 ATGTCGTTGATGTTTACATTCGTTATTTAGCTAATAAACTCAAAACCTTTTAAATAAAGAAA 2479
 QY 704 AATGATTGAACAGTTCGTTGGGTTGGGTATGATACCATGACAAACG---TAAATT 760
 DB 2478 AATCCATAGAAACAGTACGTTGGGTAGGGTATGTTGATTCGATTAAGCGCCAAAATTT 2419
 QY 761 GCSCAATAACTGGATTATTGTTACACGATGATTACGTTTGTGTCACGATATTTTGTGTTTG 820
 DB 2418 AATATATAATGATGCTTATTACGACGCTCAATACCTTCACGACATATATTCTTTCTG 2359
 QY 821 TTTAATATTATTTTCTTTGTAAGATACACTGTCATATATAGTGGTTCGATGATGATGATGATG 880
 DB 2358 CTTAATATTATTTCTTTTAAAGATATCTTTACGAAGTAGTCAAAATTTTACGAACTGA 2299
 QY 881 ACCAAGCTCAAGGATATTAATTAATTTTCAATCTTAAGCCCTGTTAAAGATATATCTGC 940
 DB 2298 AAGAAGTTCAATGATATCCCAATCTGTTCCATCTTAAATCTTTAAGTATATATCTGC 2239
 QY 941 ATTAGACTTGAATGCATCTTTTAGGTAAATTTTCAAGAGATAAATTTATTTATGATGACATAA 1000
 DB 2238 ATTGGATTTAAATGCATCTTACGAATTTTCAAGAAATATCTGATCTATGATGATACGG 2179
 QY 1001 TAATAAATTTTGAACATCGAATGATTAACACAGTGAAGTTGAACCGGTTATGAACA 1060
 DB 2178 TAGAAAGTTAAATTCAAACATCAATGATATACACTTGTCTTATGATAAACAATAATGATTT 2119
 QY 1061 CCGTTATTTTACGCGGTAATAAATAAAGCGTATAAGGCAATTAATTTTAAATTTAA 1120
 DB 2118 CAAACATCTGTAACGATATACATATCAAGAAGCCGATTAATTTACTTAGTAATAC 2059
 QY 1121 AGAACCAATTAACAACGCAAGATTTCAAAGGGTATAGCTTGTAAATTCATTCCTAGAAAA 1180
 DB 2058 TGAACCTATACGTTTGAAGATTTTCTGATACAGTATTTAGTCCATCTCTCTCAAA 1999
 QY 1181 TTATGATAACATCGTAAATTCATTTATATCATTTGCGTGGCAATTTGGAGTATGCAAC 1240
 DB 1998 TTATGATAATCTCGTTAAATCACTTTATATCGTTGCACTTGTCTTTGGATTAATGCAAC 1939
 QY 1241 AATTATAACGTCACAACTAGTTATGTTTCAACACAAATTTACTATAACCGCTTGTGAC 1300
 DB 1938 CATTTATCTGCGCGGTGAGTTATATCTTTCTTGCATAATTTACTAAACCGATGTTAC 1879
 QY 1301 TTTATCAAAATAAATGATTCGATGAGATTCGACGAGATGGTGTTCATAAATAAATTTGCAATTA 1360
 DB 1878 AATGTCCTCAATAAATGAATCAATTAAGAGAGATGGTTCATAAATAAATTTGCAATTAAC 1819
 QY 1361 TACAATTTATGAAGAAATGATTAATTTAGCAATAAGTAAATGAGATGATGATGATGATGATGAT 1420
 DB 1818 TACAATTTATGAAGAAACAGATTAATTTAATTTGATATCTTTTAAATGAAATGATGATGATGAT 1759
 QY 1421 TGAAGATCATTTAATCAACAGACAAATTTGTTGAAGTGGCTGACATGATGATGATGATGATGATGAT 1480
 DB 1758 AAGAAGATCTTTTAAATCAGCAACGTCATTTGTCGAGGATGCTTACACAGAAATTAAGAAC 1699
 QY 1481 ACCATTACAAATTTATTTCAAGGTCATTTAAATTTGATTCAGCGATGGGAAAAAAGACCC 1540
 DB 1698 GCCACTGCAGATTTTCAAGGTCATCTAAATTTAATCCACGTTGGGGGAAAAAAGATCC 1639
 QY 1541 AGCAGTTATGAAGAAATCGTTAAATTTCTATTGTAAGAAATGAATGATGATGATGATGATGATGAT 1600
 DB 1638 AGCAGTTTGGGAAGATCTTTGAAATTTTCAATTTGAAGAAGTGAATGCAATTAACAAACT 1579
 QY 1601 AGTCGAAGATTTACTTTGAATTTGATTAAGGATGTAATGATGATGATGATGATGATGATGATGATGAT 1660
 DB 1578 TGTCGAAGAACTACTCTTTTACTTACCAGAGATAGAGTCAATCATTAATGTTTGGAAATGTGA 1519
 QY 1661 GACCGTGCATATTAAATGATGAAATTCGTCGCGAAATACACTCATTAAAAACAATTTGCATCC 1720
 DB 1518 AATGTAGACGTAAATAGCGAGATTCAATCACTGCTGAGTCACTGCAACACCTACATCC 1459
 QY 1721 TGATTAATCAATTTGATACGGATCTGACATCTAAAAATCTAGAAATCTAGAAATTAATAAGAAACCTCA 1780
 DB 1458 AGATTTACTTTTGAACACATCTTCTGCTACTAAGCCTATCTCAATTAATAAATAAACCCTCA 1399


```
QY 900 AATAAATTTATTTTCAATCTTAAGCCCTGTTAAAGATATATCTGCTAGATGATTTGAATGCAATC 959
DB 954 GCCATCTGTTCCATCTTAATCTTTAAGTGATATATCTGCTAGATGATTTGAATGCAATC 895
QY 960 TTAGGTAAATTTTCAAGAGATAATTTATTTATGAGAGATATAATAAATTTATTTGAGACA 1019
DB 894 TTAGAAAATTTTCAAGAAATCTGCTATGATGATGATGATGATGATGATGATGATGATGATGAT 835
QY 1020 TCGAATGATACACAGTAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1079
DB 834 TCAATGATATAATACACTTGTCTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 775
QY 1080 ATAAAAAAGCGTATATAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1139
DB 774 CATATTCAGAGCCAGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 715
QY 1140 GATTTCAAGGGTATAGCTTTGTTAAATTCATTCAGTAGAATAATTTATGATGATGATGATGATGAT 1199
DB 714 GATTTTCTGGATACAGTGATTTAGTCCATCTCTTCAAAATTTATGATGATGATGATGATGATGAT 655
QY 1200 TCATGATATCATTTGGCTGGCATTTGGAGTGATGATGATGATGATGATGATGATGATGATGATGAT 1259
DB 654 TCATTTTATATCGTTGCACTTGTCTTTGGATTAATTTGCAACCATTTATTTACTGCGGCGTG 595
QY 1260 AGTTATGATTTTCAACACAAATTTACTAAACCGCTTCTGCTAGTTTATCAAAATAAATGATTT 1319
DB 594 AGTTATATCTTTCTCGCAATTTACTAAACCGATGATTTCAATGTCCTCAATAAATGATTT 535
QY 1320 GAGATTCGACGAGATGTTTTCAAAATAAATTTGCAATTTAAATAACAAATTTATGAGAAATA 1379
DB 534 CAAATTAAGAGAGATGTTTTCAAAATAAATTTGCAATTTAAATAACAAATTTATGAGAAATA 475
QY 1380 GATAATTTAGCAATAGCTTTTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1439
DB 474 GATAATTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 415
QY 1440 CAAAGACAAATTTGTTGAAGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1499
DB 414 CAAGCTCAATTTGCTGAGGATGCTTCAACGAATTAAGAACGCCAGCTGCAATTTATTTCAA 355
QY 1500 GGTCAATTTAAATTTGATTTACGCGATGCGGAAAAAAGAACCCAGCAGTATTAGAGATCG 1559
DB 354 GGTCAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTT 295
QY 1560 TTAATATTTTCTATTGAAGAAATCAATCGTATCATATAAATTTAGTCGAAGAAATTTACTTGAA 1619
DB 294 TTGAATATTTCAATTTGAAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 235
QY 1620 TTGACTAAAGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1679
DB 234 CTTACCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 175
QY 1680 GAAATTCGCTCGCGAATACACTCATTTAAAAAATTTGATGATGATGATGATGATGATGATGATGAT 1739
DB 174 GAGATTTCAATCACGTGAGTCACTCAACACCTCATCCAGATATATCTTTTGAACA 115
QY 1740 GATCTGACATCTAAAATCTAGAAAATTTAAATGAAACCTCATCAATTCGAACAAATTTATTT 1799
DB 114 CATCTGCTACTAGCGCTATCCCAATTTAAAAATTTAAACCGTCATCAGTTTGAACAACCTTTA 55
QY 1800 TTAATCTTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1853
DB 54 CTCAATTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1
```

RESULT 8

AAT72328/c

ID AAT72328 standard; cDNA; 736 BP.

XX AC AAT72328;

XX DT 17-OCT-2003 (revised)

```
DT 16-JAN-1998 (first entry)
XX Staphylococcus aureus cDNA encoding a novel response regulator protein.
DE Antibacterial; anti-Staphylococcus; gene therapy; bacterial infection;
XX prophylaxis; two component signal transduction system; TCSTS; Resd;
KW response regulator; Staphylococcus; Bacillus; vaccine; wound treatment;
KW dental; surgery; antibody; de.
XX Staphylococcus aureus; WCUH29.
OS Staphylococcus aureus; WCUH29.
XX Key Location/Qualifiers
FH complement(13.672)
FT CDS /tag= a
FT /product= "response_regulator_protein"
XX
XX WO9723506-A1.
XX 03-JUL-1997.
XX 20-DEC-1996; 96WO-GB003261.
XX 22-DEC-1995; 95GB-00026359.
XX (SMIK ) SMITHKLINE BEECHAM PLC.
XX Wallis N, Hodgson JE;
XX WPI; 1997-350970/32.
XX P-PSDB; AAM19274.
XX Isolated nucleic acid encoding Staphylococcus aureus response regulator -
XX useful to treat or prevent bacterial infections, in vaccines and for
XX diagnosis.
XX Claim 4; Page 45-46; 59pp; English.
XX The present sequence represents a polynucleotide, encoding a novel
XX response regulator protein from Staphylococcus aureus. DNA was isolated
XX from a chromosomal library of NCIMB 40771. The encoded polypeptide is a
XX response regulatory component of the bacterial two component signal
XX transduction system (TCSTS), related to the Resd protein of Bacillus
XX subtilis (42% identity). Inhibition of the polypeptide prevents the
XX bacteria from establishing or maintaining infection by preventing them
XX from producing the factors necessary for pathogenesis. The present DNA
XX can be used to produce recombinant polypeptides in vivo and in gene
XX therapy. The antagonist and agonists can also be used to inhibit
XX bacterial infection. Typical applications are in wound treatment (to
XX prevent bacterial adherence), prophylaxis in dental or other surgery (as
XX an alternative or adjunct to antibiotic therapy) or for treating in-
XX dwelling devices immediately before insertion. Detection of the present
XX polynucleotide or the polypeptide, by usual hybridisation or
XX immunoassays, can be used to diagnose infection. Elevated levels of
XX expression would indicate bacterial infection. The polypeptide can also
XX be used in protective vaccines to induce antibody production. The
XX antibodies themselves can be used as a diagnostic reagent or as a
XX therapeutic antagonist. (Updated on 17-OCT-2003 to standardise OS field)
XX
XX SQ Sequence 736 BP; 213 A; 138 C; 97 G; 288 T; 0 U; 0 Other;
Query Match 33.4%; Score 736; DB 2; Length 736;
Best Local Similarity 100.0%; Pred. No. 1.8e-125;
Matches 736; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 24 TAGAAAAATACAAAATTTAAAAATATATTTACCTTATTAGAAAAAGTCGTATATGAGTGT 83
DB 736 TAGAAAAATACAAAATTTAAAAATATATTTACCTTATTAGAAAAAGTCGTATATGAGTGT 677
QY 84 ACAATGACGCAATTTTAATAGTAGAGATGACAAACTTACGAGATTTCTTGATTT 143
DB 676 ACAATGACGCAATTTTAATAGTAGAGATGACAAACTTACGAGATTTCTTGATTT 617
QY 144 GGAACTCACACATGAAAAATTTACAATGTGGACACAGAGATGATGACGACGCTTTTGA 203
```

Db 616 GGAACCTCACATGAAATTAACAATGTGCACACAGAGTATGATGCACAAGACGGTTTAGA 557
Qy 204 TAAAGCGCTTAGCCATTACTATGATTTAATCATATTTAGATTTAATGTTGGCGTCAATTAA 263
Db 556 TAAAGCGCTTAGCCATTACTATGATTTAATCATATTTAGATTTAATGTTGGCGTCAATTAA 497
Qy 264 TGGCTTAGAAATTTGTCGCAAAATTTAGACAACAACAATCTACACCTATCATTTAATTAAC 323
Db 496 TGGCTTAGAAATTTGTCGCAAAATTTAGACAACAACAATCTACACCTATCATTTAATTAAC 437
Qy 324 AGCGAAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 383
Db 436 AGCGAAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 377
Qy 384 AGTTAAGCGCTTAGCCATTACTATGATTTAATCATATTTAGATTTAATGTTGGCGTCA 443
Db 376 AGTTAAGCGCTTAGCCATTACTATGATTTAATCATATTTAGATTTAATGTTGGCGTCA 317
Qy 444 GCCACAAAGGATATTTATCGATGTCACCGGTATTAACAATTTGATAAGAACGGCTTTTAAAGT 503
Db 316 GCCACAAAGGATATTTATCGATGTCACCGGTATTAACAATTTGATAAGAACGGCTTTTAAAGT 257
Qy 504 GACGTAATGCGCGAGAAATTTGAATTAACAACAACAAGATGATGATGATGATGATGATGATGAT 563
Db 256 GACGTAATGCGCGAGAAATTTGAATTAACAACAACAAGATGATGATGATGATGATGATGATGAT 197
Qy 564 AGCTGAAATTAACCAATTTGATGCAACGGGAAACAAATTTTAAATCATGATGATGATGATGAT 623
Db 196 AGCTGAAATTAACCAATTTGATGCAACGGGAAACAAATTTTAAATCATGATGATGATGATGAT 137
Qy 624 TAATAGTGAAGTAGAACAACAATGTCGTAGATGTTTATATAGATATTTACGAAACAAGTT 683
Db 136 TAATAGTGAAGTAGAACAACAATGTCGTAGATGTTTATATAGATATTTACGAAACAAGTT 77
Qy 684 AAAACCATACGATCTGACAAATGATTTGAACAGCTTGTGGCGTGGGTATGATGATGATGAT 743
Db 76 AAAACCATACGATCTGACAAATGATTTGAACAGCTTGTGGCGTGGGTATGATGATGATGATGAT 17
Qy 744 ATGACAAACGTAAT 759
Db 16 ATGACAAACGTAAT 1

RESULT 9
ID AAX90602/c
XX AAX90602 standard; DNA; 736 BP.
XX AC AAX90602;
XX DT 01-OCT-1999 (first entry)
XX DE Cognate response regulator, polynucleotide, of Histidine kinase.
XX KW Cognate response regulator; Two component signal transduction system;
KW TCSTS; Histidine kinase; virulence; bacterial pathogenesis; screening;
KW autophosphorylate; antibacterial agent; Helicobacter pylori; ds.
XX OS Staphylococcus aureus.
XX FH Location/Qualifiers
CDS complement(13..672)
FT /*tag= a
FT /product= "Cognate response regulator to histidine
FT kinase"
FT /note= "shows 42% identity to ResD response regulator
FT protein from Bacillus subtilis"
XX PN W0936508-A1.
XX PD 22-JUL-1999.
XX PF 12-JAN-1999; 99WO-0000610.

XX 13-JAN-1998; 98US-00006627.
XX (SMIK) SMITHKLINE BEECHAM CORP.
XX Wallis NG, Shilling LK, Mooney JL, Debouck C, Zhong Y;
PI Jaworski DD, Wang M, Throup JP;
XX WPI; 1999-444390/37.
XX P-PSDB; AAY28602.
XX Novel histidine kinase polynucleotides and polypeptides used to screen
PT for antibacterial compounds.
XX PS Claim 24; Page 7; 43pp; English.
XX The present polynucleotide sequence encodes the Staphylococcus aureus
CC response regulator cognate to histidine kinase. Histidine kinase is a
CC component of the two component signal transduction systems (TCSTS) which
CC is involved in bacterial pathogenesis and virulence. It undergoes
CC autophosphorylation at a histidine residue and the phosphate group is
CC transferred to the cognate response regulator. This novel response
CC regulator shows 42% identity to the ResD response regulator protein from
CC Bacillus subtilis. The cognate response regulator can be used for
CC screening drugs which interfere with the interaction of histidine kinase
CC and the response regulator. The histidine kinase products can be used to
CC screen and identify new antibacterial drugs, agonists and antagonists
CC effective against Helicobacter pylori infections and infections caused by
CC resistant S. aureus strains
XX SQ Sequence 736 BP; 213 A; 138 C; 97 G; 288 T; 0 U; 0 Other;
Query Match 33.4%; Score 736; DB 25; Length 736;
Best Local Similarity 100.0%; Pred. No. 1.8e-125;
Matches 736; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 24 TAGAAAAATACAAAATTTTAAATATATTTACCTTATTAGAAAAAGTCGTAATATGAGGTG 83
Db 736 TAGAAAAATACAAAATTTTAAATATATTTACCTTATTAGAAAAAGTCGTAATATGAGGTG 677
Qy 84 ACAATGACGCAAAATTTTAAATAGTAGAGATGACAAACAACTTACGAGATTTCTTGAATT 143
Db 676 ACAATGACGCAAAATTTTAAATAGTAGAGATGACAAACAACTTACGAGATTTCTTGAATT 617
Qy 144 GGAATCTCACATGAAATTTAACAATGTGGACACAGAGTATGATGCACAAGACGGTTTAGA 203
Db 616 GGAATCTCACATGAAATTTAACAATGTGGACACAGAGTATGATGCACAAGACGGTTTAGA 557
Qy 204 TAAAGCGCTTAGCCATTACTATGATTTAATCATATTTAGATTTAATGTTGGCGTCAATTAA 263
Db 556 TAAAGCGCTTAGCCATTACTATGATTTAATCATATTTAGATTTAATGTTGGCGTCAATTAA 497
Qy 264 TGGCTTAGAAATTTGTCGCAAAATTTAGACAACAACAATCTACACCTATCATTTAATTAAC 323
Db 496 TGGCTTAGAAATTTGTCGCAAAATTTAGACAACAACAATCTACACCTATCATTTAATTAAC 437
Qy 324 AGCGAAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 383
Db 436 AGCGAAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 377
Qy 384 AGTTAAGCGCTTAGCCATTACTATGATTTAATCATATTTAGATTTAATGTTGGCGTCA 443
Db 376 AGTTAAGCGCTTAGCCATTACTATGATTTAATCATATTTAGATTTAATGTTGGCGTCA 317
Qy 444 GCCACAAAGGATATTTATCGATGTCACCGGTATTAACAATTTGATAAGAACGGCTTTTAAAGT 503
Db 316 GCCACAAAGGATATTTATCGATGTCACCGGTATTAACAATTTGATAAGAACGGCTTTTAAAGT 257
Qy 504 GACGTAATGCGCGAGAAATTTGAATTAACAACAACAAGATGATGATGATGATGATGATGATGAT 563
Db 256 GACGTAATGCGCGAGAAATTTGAATTAACAACAACAAGATGATGATGATGATGATGATGATGAT 197
Qy 564 AGCTGAAATTAACCAATTTGATGCAACGGGAAACAAATTTTAAATCATGATGATGATGATGAT 623

```

196 AGCTGAAATAAAACCAATGTTATGCAACGCGAACAATTTTAAATCATGTATGGGTTA 137
136 TAATAGTGAAGTAGAACAATAATGTCGTAGATGTTTATATAGATATTTACGAAACAAGTT 683
136 TAATAGTGAAGTAGAACAATAATGTCGTAGATGTTTATATAGATATTTACGAAACAAGTT 77
684 AAAACCATACATCGTGACAAAATGATTGAAACACAGTTTCGTGGCGTTGGGTATGTGATACG 743
76 AAAACCATACATCGTGACAAAATGATTGAAACACAGTTTCGTGGCGTTGGGTATGTGATACG 17
744 ATGACAAAACGTAAT 759
16 ATGACAAAACGTAAT 1

RESULT 10
ABN90850
ID ABN90850 standard; DNA; 1416 BP.
XX
AC ABN90850;
XX
DT 24-JUL-2002 (first entry)
XX
DE Staphylococcus epidermidis ORF nucleic acid sequence SEQ ID NO:313.
XX
KW Staphylococcus epidermidis; open reading frame; ORF; bacterial infection;
KW antibacterial; gene therapy; gene; ds.
XX
OS Staphylococcus epidermidis.
XX
PN US6380370-B1.
XX
PD 30-APR-2002.
XX
PF 13-AUG-1998; 98US-00134001.
XX
PR 14-AUG-1997; 97US-0055779P.
XX
PR 08-NOV-1997; 97US-0064964P.
XX
XX (GENO-) GENOME THERAPEUTICS CORP.
XX
XX Doucette-Stamm LA, Bush D;
XX
XX WPI; 2002-381255/41.
XX
XX P-PSDB; ABP38305.
XX
XX Novel isolated nucleic acid encoding a Staphylococcus epidermis
XX polypeptide, useful for diagnosing and treating bacterial infections.
XX
XX Disclosure; SEQ ID NO 313; 267pp; English.
XX
XX ABN90538 to ABN93374 represent Staphylococcus epidermidis open reading
XX frame (ORF) nucleic acid sequences which encode the amino acid sequences
XX given in ABP35124 to ABP37960. The S. epidermidis sequences have
XX antibacterial activity and can be used in gene therapy. The sequences can
XX also be used in the diagnosis and treatment of bacterial infections,
XX particularly S. epidermidis infections. The sequences can be used to
XX screen for compounds able to interfere with the S. epidermidis life cycle
XX or inhibit S. epidermidis infection. N.B. The sequence data for this
XX patent did not form part of the printed specification, but was obtained
XX in electronic format directly from the USPTO web site
XX
XX Sequence 1416 BP; 529 A; 216 C; 223 G; 448 T; 0 U; 0 Other;
XX
XX Query Match 33.3%; Score 732.8; DB 6; Length 1416;
XX Best Local Similarity 70.7%; Pred. No. 7.5e-125;
XX Matches 990; Conservative 0; Mismatches 407; Indels 3; Gaps 1;
XX
XX 700 GACAAAATGATTGAAACAGTTCGTGGCGTTGGGTATGTGATACGATGACAAAACGT---A 756
XX 2 GAAAAATCCATAGAAACAGTACGTCGTGGCGTTGGGTATGTGATTCGATGATTAAAGCGCAAA 61

```



```
ACF74591
ID ACF74591 standard; DNA; 657 BP.
AC ACF74591;
XX
XX 20-NOV-2003 (first entry)
XX
XX Staphylococcus aureus DNA #2271.
XX
XX Antibacterial; vaccine; gene therapy; infection; sepsis; diagnosis;
XX enzymatic assay; antibiotic target; gene; ds.
XX
XX Staphylococcus aureus.
XX
XX WO200294868-A2.
XX
XX 28-NOV-2002.
XX
XX 27-MAR-2002; 2002WO-IB002637.
XX
XX 27-MAR-2001; 2001GB-00007661.
XX
XX (CHIR-) CHIRON SPA.
XX
XX Maignani V, Mora M, Scarselli M;
XX
XX WPI: 2003-120786/11.
XX P-PSDB; ABM73031.
XX
XX New Staphylococcus aureus protein, useful as a vaccine for treating or
XX preventing Staphylococcal infection, specifically an infection caused by
XX S. aureus, e.g. sepsis.
XX
XX Claim 6; SEQ ID NO 4541; 49pp; English.
XX
XX The invention relates to novel genes and encoded proteins from
XX Staphylococcus aureus. A composition comprising the S. aureus protein, a
XX nucleic acid encoding the protein, or an antibody to the protein, is
XX useful as a pharmaceutical, particularly as a vaccine for treating or
XX preventing infection due to Staphylococcus bacteria, specifically an
XX infection caused by S. aureus. The composition is particularly useful for
XX treating or preventing sepsis in a patient. The composition can also be
XX used for diagnostics. The protein is also used in an assay for enzymatic
XX studies and as a target for antibiotics. This sequence represents one of
XX the novel S. aureus genes of the invention
XX
XX Sequence 657 BP; 248 A; 90 C; 128 G; 191 T; 0 U; 0 Other;
XX
XX Query Match 29.9%; Score 657; DB 7; Length 657;
XX Best Local Similarity 100.0%; Pred. No. 5e-111;
XX Matches 657; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 88 ATGACGCAATTTTAAATAGTAGAAGATGAACAAACTTAGCAAGATTTCTTGAATTGGAA 147
XX |
XX 1 ATGACGCAATTTTAAATAGTAGAAGATGAACAAACTTAGCAAGATTTCTTGAATTGGAA 60
XX |
XX 148 CTCACACATCAATTAATGATGAGACACAGATGATGACAGACGCTTTAGATAAA 207
XX |
XX 61 CTCACACATCAATTAATGATGAGACACAGATGATGACAGACGCTTTAGATAAA 120
XX |
XX 208 GCGCTTAGCCATTACTATGATTTAATCATATGATTTAATGTCGCTCAATTAATGGC 267
XX |
XX 121 GCGCTTAGCCATTACTATGATTTAATCATATGATTTAATGTCGCTCAATTAATGGC 180
XX |
XX 268 TTAGAAATTTGCGCAAAATTAGACAAACAATCTACCTATCATTAATTAATACAGCG 327
XX |
XX 181 TTAGAAATTTGCGCAAAATTAGACAAACAATCTACCTATCATTAATTAATACAGCG 240
XX |
XX 328 ABAAGTATACGATGACAAAGTGTGGCTGATTTACGTCGACGAGTATATAGTT 387
XX |
XX 241 ABAAGTATACGATGACAAAGTGTGGCTGATTTACGTCGACGAGTATATAGTT 300
XX |
XX 388 AAGCCGTTTGATATTGAAGAACTTTTAGCAAGAATTCGTGCAATTTTACGTCGTCAGCCA 447
XX |

Db 301 AACCGTTTGTATATGAGAAGCTTTTAGCAAGAAATTCGTGCAATTTTACGTCGTCAGCCA 360
Qy 448 CAAAAGGATATATCGATGTCACCGGTATTAACAATTGATAAGAACGCTTTTAAAGTAGCG 507
Db 361 CAAAAGGATATATCGATGTCACCGGTATTAACAATTGATAAGAACGCTTTTAAAGTAGCG 420
Qy 508 GTAAATGGCGCAGAAATTTGAATTAACAAAACAGAGTATGATTTACTATATCTTAGCT 567
Db 421 GTAAATGGCGCAGAAATTTGAATTAACAAAACAGAGTATGATTTACTATATCTTAGCT 480
Qy 568 GAAATATAAAACCATGTTATGCAACCGGAAACAAATTTTAAATCATGTATGGGTATAAT 627
Db 481 GAAATATAAAACCATGTTATGCAACCGGAAACAAATTTTAAATCATGTATGGGTATAAT 540
Qy 628 AGTGAAGTAGAACAACAAATGTCGTAGATGTTTATATAAGATATTTAGCAACAAGATTAAAA 687
Db 541 AGTGAAGTAGAACAACAAATGTCGTAGATGTTTATATAAGATATTTAGCAACAAGATTAAAA 600
Qy 688 CCATACGATCGTGACAAATGATTGAAACAGTTTCGTGGCGTTGGGTATGTGATACGA 744
Db 601 CCATACGATCGTGACAAATGATTGAAACAGTTTCGTGGCGTTGGGTATGTGATACGA 657

RESULT 13
AAH53016
ID AAH53016 standard; DNA; 855 BP.
XX
XX AAH53016;
AC
XX 03-SEP-2001 (first entry)
XX
XX S. epidermidis open reading frame nucleotide sequence SEQ ID NO:1425.
XX
XX Staphylococcus epidermidis SR1 strain; infection; diagnosis; vaccination;
XX endocarditis; ds.
XX
XX Staphylococcus epidermidis.
XX
XX WO200134809-A2.
XX
XX 17-MAY-2001.
XX
XX 09-NOV-2000; 2000WO-US030782.
XX
XX 09-NOV-1999; 99US-0164258P.
XX
XX (GLAX ) GLAXO GROUP LTD.
XX
XX Kimmerly WJ;
XX
XX WPI; 2001-316495/33.
XX P-PSDB; AAG82166.
XX
XX Nucleic acids encoding polypeptides from Staphylococcus epidermidis,
XX useful for vaccinating against infections, e.g. endocarditis.
XX
XX Claim 8; Page 402-403; 2188pp; English.
XX
XX AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides
XX (II), given in AAG81454 to AAG83120, from Staphylococcus epidermidis. (I)
XX and (II) can have antibacterial activity and therefore can be used in
XX vaccination. The nucleic acids (I) may be used to produce the S.
XX epidermidis polypeptides (II) via the production of vectors containing
XX them which are used to produce hosts cells which express the
XX polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
XX used to vaccinate subjects and to raise antibodies against the bacteria.
XX The polypeptides may also be used to assay for other inhibitors of their
XX activity and therefore identify compounds that may be used for the
XX treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to
XX AAH55090 represent specifically claimed S. epidermidis genomic DNA
XX polynucleotide sequences from the present invention. AAH55091 to AAH55098
XX represent oligonucleotide sequences and primers which are used in the
```


CC exemplification of the present invention. N.B. The present invention
CC specifically claims all the polynucleotide sequences given in the
CC sequence listing of the present specification, however the sequence
CC listing only goes up to SEQ ID NO:4454 so even though sequences are given
CC in the disclosure for SEQ ID NO:4465 to 4472, no sequences are present
CC for SEQ ID NO:4455 to 4464

SQ Sequence 855 BP; 332 A; 130 C; 137 G; 256 T; 0 U; 0 Other;

Query Match 21.4%; Score 471; DB 4; Length 855;
Best Local Similarity 72.6%; Pred. No. 5.5e-77;
Matches 609; Conservative 0; Mismatches 230; Indels

Qy	1258	TCGTTATGTA	TTTTCAACACAA	TAATTA	CTAAACCGCTTG	TCAGTTTATCAAA	TAATAAATGA	1317
D5 <td>2 <td>TGAGTTATAT <th>CTTTTCTTCGCA</th> <th>AAATTA</th> <th>CTAAACCGAGAT</th> <th>TTACAAATGTC</th> <th>CAATTAATAAATGA</th> <th>61</th> </td></td>	2 <td>TGAGTTATAT <th>CTTTTCTTCGCA</th> <th>AAATTA</th> <th>CTAAACCGAGAT</th> <th>TTACAAATGTC</th> <th>CAATTAATAAATGA</th> <th>61</th> </td>	TGAGTTATAT <th>CTTTTCTTCGCA</th> <th>AAATTA</th> <th>CTAAACCGAGAT</th> <th>TTACAAATGTC</th> <th>CAATTAATAAATGA</th> <th>61</th>	CTTTTCTTCGCA	AAATTA	CTAAACCGAGAT	TTACAAATGTC	CAATTAATAAATGA	61

Accession	Sequence	Length
Qy	TTGAGATTTCGACGAGATGGTTTTTCAAAATAAAATTCGAATTAAATACAAATTATGAAGAA	1377
Db	ATCAGATTTTGAAGAGATGGTTTTTCAAAATAAAATTCGAATTAAATACAAATTATGAAGAA	121

Qy		1378	TGATATAATTAGCAAAATACGTTTAAATGAGATGAGTAGGCCAAATTAAGAAGATCATTTTAATC	1437
Db		122	CAGATAATTTAATGTGAATCTTTTAAATGAAAATGATGTATCAATATAGAAGATCTTTTAATC	181

Qy 1438 AACAAAGCAATTTGTTGAAGATCGTCACATGAATTACGACACCCATTACAAATTATTC 1497

Dd 182 AGCAACGTCATTTGTGTCAGGATGCTTTCACACGAATTAAGACGCCACTGCGATTTATTC 241

Qy	1498	AAGGTCATTAAATTTGATTCAGCGATGGGAAAAAGACCCAGCAGTATTAGAGAAAT	1557
Dh	242	AGGTCATCTAATTTTATTCGACCTGGGGCCGNNNNNNAGCTGACCGCTTTTCCGNNAT	291

Qy 1558 CGTAAATATTTCTATTGAAGAAATGAATCGTATCATAAAAATTAGTCGAAGAAATTACTTG 1617

QY 1618 AATTGACTAAAGGAGATGTAATAATGACATTTCTTCTGAAGCGGAGACCGTGCATATTAAATG 1677

QY 1678 ATGAAATTCGCTCGGAATACACTCATTAAACAATTCATCTGATTATCAATTTGATA 1737

QY 1738 CGGATCTGACATCTTAAAAATCTAGAAATTAAATGAAACCTCATCAATTCGAAACAATTAT 1797

Q_Y 1798 TTTTAACTTTTATTGATAATGCAATCAAATATGATGTCAGAAGTAAGAAAAATTAAAGTTA 1857

Q7 1858 AGCAAGGTTAAAAAATNAGCAAAAATAATTCAAAATTCAGATCATGGAATTGGTATTC 1917

Qy 1918 CACAGGAAGATCAAGATTTCATTTTTCGCTTTTATCGAGTGGATAAATCTCGTTCAA 1977

Qy 1978 GAAGTCAAGGGCGGTAAATGGACTCGGATTATCTATTGCTCAAAAAATCATCTCAATTAACG 2037

Qy 2038 GAGGATCGATTAAATAATTAAAGTGGAATTAATAAGGACAACGTTTTAAAAATCATATTT 2096

702 GAGGATGCTTCCAGGTGAGAGAGTGGACCTACGAGGATGACGACCTTCATATTCAGTCTT 830

AAH53698
ID AAH53698 standard; DNA; 855 BP.

XX
AC
XX
XX
DT
XX

AAH53698;
03-SEP-2001 (first entry)
S. epidermidis open reading frame nucleotide sequence SEQ ID NO:2789.

Staphylococcus epidermidis SR1 strain; infection; diagnosis; vaccination;

emboacaulitis; us.
XX
OS
XX
Staphylococcus epidermidis.
XX

17-MAY-2001.

09-NOV-2000; 2000WC-05030782.
XX
XX
09-NOV-1999; 99US-0164258P.
XX
XX

FR
XX
XX
PI
XX
XX

(GERRA / GERRA GROUP LTD.
Kimmerly WJ;

XX Nucleic acids encoding polymyxins from *Staphylococcus epidermidis*
DR P-PSDB; AAG82848.
DR MF1; 2001-316435/33.
DR

xx useful for vaccinating against infections, e.g. endocarditis.
 PT
 PS
 yy Claim 8; Page 725-736; 2188pp; English.

CC AAHS2304 to AAHS3970 represent nucleic acids (I) encoding polypeptides
CC (II), given in AA81454 to AA83120, from *Staphylococcus epidermidis*. (I)
CC (II), given in AA81454 to AA83120, from *Staphylococcus epidermidis*. (I)
CC CC and (II) can have antibacterial activity and therefore can be used in
C

epidermidis polypeptides (II) via the production of vectors containing them which are used to produce hosts cells which express the polypeptides. The polypeptides (II) (and/or nucleic acids) may then be

CC The polypeptides may also be used to assay for other inhibitors of their
CC activity and therefore identify compounds that may be used for the
CC treatment of *S. epidermidis* infections, e.g. endocarditis. AAH53971 to

CC polynucleotide sequences from the present invention. AAH55091 to AAH55098
CC represent oligonucleotide sequences and primers which are used in the
CC exemplification of the present invention. N.B. The present invention

CC sequence listing of the present specification, however the sequence
CC listing only goes up to SEQ ID NO:4454 so even though sequences are given
CC in the disclosure for SEQ ID NO:4465 to 4472, no sequences are present

xx
SQ Sequence 855 BP; 332 A; 130 C; 137 G; 256 T; 0 U; 0 Other;

Best Local Similarity 72.6%; Pred. No. 5.5e-77; Mismatches 0; Gaps 0;
Matches 609; Conservative 230; Indels 0

2 TGAGTTTATATCTTTCTCGCAAAATTACTAAACCGATAGTTTCAATGTCCAATATAAATGA 61

D_b 62 ATCAAATTAGAAGAGATGGTTTTCARATTAACCTTGAATTAACTACAAATTATGAGA 121

Qy 1378 TAGATTAATTTAGCAAAATACGTTTTAAATGAGATGATGAGCCAAAATGGAAGAATCATTTAAATC 1435

Db 122 CAGATAATTTAATGTGATCTTTTAAATGAAATGATGTATCAAATAGAAGAATCTTTTAAATC 181

QY
1438 AA CAAAGACAATTTGTTGAAGATCGGTACCATGAATTACGAAACCACTTACAAATTATTC

Db 182 AGCAACGTCATTTGTCGAGATGCTTACACAGAAATTAAGACGCCACTGCAGATTATTC 241
Qy 1498 AAGTCATTTAAATTTGATTCAGGATGCGGAAAGAACCCAGAGTATTAGAAGAT 1557
Db 242 AAGTCATCTAAATTTAAATCCACGTTGGGGGAAAAAGATCCAGCAGTTTGGAGAAAT 301
Qy 1558 CGTTAAATATTTCTATTGAAGAAATGAATCGTATCATATAAAATAGTGCAGAAATTAATTG 1617
Db 302 CTTTGAATATTTCAATTTGAAGAGTGAATCGAATACAAAATCTGTGCAAGAACTACTTT 361
Qy 1618 AATTGACTAAAGAGATGTAATGACATTTCTTCTGAAGCGCAGACCGTGCATATTAATG 1677
Db 362 TACTTACCAAGATAGAGTCAATCATATGTTTGGAAATGTGAATGTAGACGTAATA 421
Qy 1678 ATGAATTCCTCGGAAATACATCTATTAACAAATTTGCAATCTGATTATCAATTTGATA 1737
Db 422 GCAGAGATTCATCACTGCTGGAAGTCACTGCAACACCTTACATCCAGATTATATCTTTGAAA 481
Qy 1738 CGATCTGACATCTAAAAATCTAGAAATTAATAATGAAACCTCATCAATTCGAACAATAT 1797
Db 482 CACATCTTGCTACTAGCCCTATCCAAATTAATAATTAACCGTCACTAGTTTGAACAACTCT 541
Qy 1798 TTTTAAATCTTTATGTAATGCAATCAATATGATGTGAAGAAATAAGAAAATTAAGTTA 1857
Db 542 TACTCATATTTATGTAATGCAATGAAATACGACACTGAAACATAGCACATTAATAATTG 601
Qy 1858 AGACAGGTTAAAAATAAGCAAAAAATAATGAAATTCAGATCATGGAATTTGGTATTC 1917
Db 602 TTACTCAACTAAAAATAATGAATGATGATGATGATGATGATGATGATGATGATGATGATG 661
Qy 1918 CAGAGGAAGTCAAGATTTCAATTTTTCATCGCTTTTATCGAGTGGATTAATCTCGTTCAA 1977
Db 662 CAAAAGCTGACTAGAAATTTATCTTTGATGATTTTATCTGTAGATAAATCACTGCTC 721
Qy 1978 GAAGTCAAGCGGTAATGAGTCTCGGATTTATCTATGCTCAAAAAATCAATCAATTAACG 2037
Db 722 GTAGTCAAGGAGGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 781
Qy 2038 GAGATCGATTAATTAATTAAGTGAATTAATTAAGGAAACAAAGTTTAAATCATATTT 2096
Db 782 GTGGTATGATCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 840

RESULT 15

ABN90912
ID ABN90912 standard; DNA; 669 BP.
XX
AC ABN90912;
XX
DT 24-JUL-2002 (first entry)
XX
DE Staphylococcus epidermidis ORF nucleic acid sequence SEQ ID NO:375.
XX
KW Staphylococcus epidermidis; open reading frame; ORF; bacterial infection;
KW antibacterial; gene therapy; gene; ds.
XX
OS Staphylococcus epidermidis.
XX
PN US6380370-B1.
XX
PD 30-APR-2002.
XX
PF 13-AUG-1998; 98US-00134001.
XX
PR 14-AUG-1997; 97US-0055779P.
PR 08-NOV-1997; 97US-0064964P.
XX
PA (GENO-) GENOME THERAPEUTICS CORP.
XX
PI Doucette-Stamm LA, Bush D;
XX
DR WPI; 2002-381255/41.
DR P-PSDB; ABP38367.

XX
PT Novel isolated nucleic acid encoding a Staphylococcus epidermis
PT polypeptide, useful for diagnosing and treating bacterial infections.
XX
PS Disclosure; SEQ ID NO 375; 267pp; English.
XX
CC ABN90538 to ABN93374 represent Staphylococcus epidermidis open reading
CC frame (ORF) nucleic acid sequences which encode the amino acid sequences
CC given in ABP35124 to ABP37960. The S. epidermidis sequences have
CC antibacterial activity and can be used in gene therapy. The sequences can
CC also be used in the diagnosis and treatment of bacterial infections,
CC particularly S. epidermidis infections. The sequences can be used to
CC screen for compounds able to interfere with the S. epidermidis life cycle
CC or inhibit S. epidermidis infection. N.B. The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from the USPTO web site
XX
SQ Sequence 669 BP; 248 A; 89 C; 130 G; 202 T; 0 U; 0 Other;

Query Match 18.3%; Score 401.8; DB 6; Length 669;
Best Local Similarity 75.0%; Pred. No. 2.4e-54;
Matches 502; Conservative 0; Mismatches 167; Indels 0; Gaps 0;
Qy 79 GGTGTACAAATGACGCAAAATTTTAAATAGTAGAAGATGAACAACTTAGCAAGATTCTT 138
Db 1 GGTGCAATATGCAAAATATTTTAAATGTTAGAGATGACAAATCTTCTAGATTATTA 60
Qy 139 GAATGGAACTCACACATGAAAATTAACAATGGGACACAGATGATGGAACAAGCGT 198
Db 61 GAGCTTGAGTTAACTCATGAAAATTTATCTGTTGATATAGAGAAATGATGGAAGGTGGT 120
Qy 199 TTAGTAAAGCGCTTAGCCATTACTGATGATTTAAATCATATATTAGATTATGTTCCGCTCA 258
Db 121 TTGGATAAAGCAATATCTAAGCCCTATGATTTATATATTAGACTTAAGCTTCCAAAC 180
Qy 259 ATTAATGGCTTAGAAAATTTGTCGCAAAATTTAGACAAACAACTCTACACTATCATTATA 318
Db 181 ATAAATGGCTAGAAATTTGTAGACAAATTCGTCAAAAACAACTACTCCAATTATCAIC 240
Qy 319 ATTACAGCAAAAGTGATACGATGACAAAGTTGCTGGCTTGATTTACGGTGCAGACGAT 378
Db 241 ATTACTCGCAAAAGCGAGACATATGATAAAGTAGCTGGATTGGACTATCGGGCAGATGAC 300
Qy 379 TATATAGTTAAGCGCTTTCGATATTTAGAGAACTTTTGGCAAGAAATTCGTCAATTTTACGT 438
Db 301 TACATTTGAAAAACCCCTTTGATATAGAGAAATTTGTTGCAAGAAATAAGACGGTATTGCGC 360
Qy 439 CGTCAGCCACAAAAGGATATTATCGATGTCAACGGTATTACAAATTGATTAAGAACGCTTTT 498
Db 361 AGACAGCCAGATTAAGATGTTTTAGATATCAATGGTATTATTCATTGATAAGATGCTTT 420
Qy 499 AAAGTGACCGTAAATGGCGCAGAAATTTGAATTAACAAAACAGAGATGATGATTACTATAT 558
Db 421 AAAGTTACTGTTAAATGGCCATCAATAGAAATTAACCTTAAACAGAAATACGATTTACTATAT 480
Qy 559 CTTCTAGCTGAAAATAAAACCATTGTTATGCAACGGGAAACAAATTTTAAATCATGTATGG 618
Db 481 GTTTAGCTGAAAATTCGTAACCCAGTCATGCAACGGTGAACAAATTCCTCATCGATGG 540
Qy 619 GGTTAATATAGTCAAGTAGAAAACAAATGTCGTAGATGTTTATATAGATATTATTCGAAAC 678
Db 541 GGGTATAATAGTGAAGTAGAAGCAATGTCGTTGATGTTTATATTCGTTATTATACGTAAT 600
Qy 679 AAGTTAAACCATACGATCGTGACAAATGATGAAACAGTTCGGCGCTTGGGTATGTC 738
Db 601 AAACCTCAACCTTTTAAATAAGAAAAATCCATAGAAACAGTACGTCGGCGGTATGTC 660
Qy 739 ATACGATGA 747
Db 661 ATTCGATGA 669

Search completed: October 5, 2004, 03:49:24

Job time : 912.032 secs



GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 5, 2004, 03:08:42 ; Search time 166.368 seconds
(without alignments)
7341.850 Million cell updates/sec

Title: US-09-006-627-1
Perfect score: 2201
Sequence: 1 TAATTAAAGCACTATT.....AAAACAAATAACAGTGCTT 2201

Scoring table: IDENTITY_NUC
Gapop 10_0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq: *
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq: *
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq: *
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq: *
5: /cgn2_6/ptodata/2/ina/PTUS_COMB.seq: *
6: /cgn2_6/ptodata/2/ina/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2201	100.0	2201	4	US-09-642-000-1
2	1432	65.1	1440	4	US-08-956-171E-725
3	736	33.4	736	3	US-08-771-110-1
4	736	33.4	736	4	US-09-642-000-3
5	732.8	33.3	1416	4	US-09-134-001C-313
6	673.2	30.6	4858	4	US-08-956-171E-379
7	401.8	18.3	869	4	US-09-134-001C-375
8	206.2	9.4	1566	4	US-09-107-532A-2777
9	205.8	9.4	1128	4	US-09-134-000C-557
10	187.2	8.5	687	4	US-09-107-532A-2882
11	149.8	6.8	741	4	US-09-134-001C-338
12	145.8	6.6	555	4	US-09-134-000C-415
13	144	6.5	684	4	US-09-328-352-4049
14	130.2	5.9	760	4	US-08-956-171E-822
15	129	5.9	696	4	US-09-134-000C-3259
16	127.6	5.8	9425	4	US-08-956-171E-87
17	127.4	5.8	900	2	US-08-771-455-1
18	126	5.7	3731	4	US-09-082-077-1
19	125.8	5.7	696	4	US-09-134-000C-1740
20	125.6	5.7	702	4	US-09-082-077-2
21	125.2	5.7	1704	4	US-09-134-001C-925
22	125	5.7	705	4	US-09-134-001C-226
23	123.6	5.6	2244	3	US-09-272-414-1
24	121.8	5.5	587	4	US-09-134-001C-1246
25	121.2	5.5	724	3	US-08-771-321-1
26	120.8	5.5	711	4	US-09-107-532A-3219
27	119.6	5.4	3636	4	US-08-961-527-78

28	117.8	5.4	277	4	US-08-956-171E-1042
29	116.4	5.3	1827	4	US-09-082-077-3
30	116.4	5.3	2400	3	US-08-963-901-1
31	116.4	5.3	2400	3	US-08-963-901-5
32	114.8	5.2	1434	4	US-09-107-532A-1794
33	114.8	5.2	1827	3	US-09-310-275-1
34	113.8	5.2	738	4	US-08-134-001C-942
35	113.8	5.2	850	2	US-08-771-783-1
36	113.4	5.2	741	4	US-09-107-532A-481
37	112.6	5.1	9909	4	US-08-961-527-12
38	111.4	5.1	1830121	4	US-09-557-884-1
39	111.4	5.1	1830121	4	US-09-643-990A-1
40	111	5.0	4279	4	US-08-993-825-1
41	109.8	5.0	720	4	US-09-711-164-196
42	109.6	5.0	1060	1	US-08-850-118-1
43	109.6	5.0	1060	2	US-09-008-253-1
44	109.6	5.0	1060	3	US-09-093-335-1
45	109.2	5.0	1776	4	US-09-134-001C-391

ALIGNMENTS

RESULT 1
US-09-642-000-1
; Sequence 1, Application US/09642000
; Patent No. 6548281
; GENERAL INFORMATION:
; APPLICANT: Wallis, Nicola G.
; Shilling, Lisa K.
; Mooney, Jeffrey L.
; Debouck, Christine
; Zhong, YiYi
; Jaworski, Deborah D.
; Wang, Min
; Throup, John P.
; TITLE OF INVENTION: Histidine Kinase
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dechert, Price & Rhoads
; STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103-2793
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/642,000
; FILING DATE: 18-Aug-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/006,627
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Falk, Stephen T
; REGISTRATION NUMBER: 36,795
; REFERENCE/DOCKET NUMBER: GM10127
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-994-2488
; TELEFAX: 215-994-2222
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2201 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-642-000-1

Sequence 1042, Ap
Sequence 3, Appli
Sequence 1, Appli
Sequence 5, Appli
Sequence 1794, Ap
Sequence 1, Appli
Sequence 942, App
Sequence 1, Appli
Sequence 481, App
Sequence 12, Appli
Sequence 1, Appli
Sequence 1, Appli
Sequence 196, App
Sequence 1, Appli
Sequence 1, Appli
Sequence 1, Appli
Sequence 391, App

Db 2101 CATGCTGAGAGCTCAATCAAAAGTCATAGGATCAATTTTAAAGTACACATTAGCTGTGA 2160
 Qy 2161 CTAATGTATAAGAACAACTATAAAACAAATAAAACAGTGTGTT 2201
 Db 2161 CTAATGTATAAGAACAACTATAAAACAAATAAAACAGTGTGTT 2201

RESULT 2

US-08-956-171E-725
 ; Sequence 725: Application US/08956171E
 ; Patent No. 6593114
 ; GENERAL INFORMATION:
 ; APPLICANT: Charles Kunsch
 ; Gil H. Choi
 ; Patrick S. Dillon
 ; Craig A. Rosen
 ; Steven C. Barash
 ; Michael R. Fannon
 ; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
 ; NUMBER OF SEQUENCES: 5256
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Human Genome Sciences, Inc.
 ; STREET: 9410 Key West Avenue
 ; CITY: Rockville
 ; STATE: Maryland
 ; COUNTRY: USA
 ; ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
 COMPUTER: HP Vectra 486/33
 OPERATING SYSTEM: MSDOS version 6.2
 SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/956,171E
 FILING DATE: 20-Oct-1997
 CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/009,861
 FILING DATE: January 5, 1996
 APPLICATION NUMBER: 08/781,986
 FILING DATE: January 3, 1997

ATTORNEY/AGENT INFORMATION:

NAME: Mark J. Hyman
 REGISTRATION NUMBER: 46,789
 REFERENCE/DOCKET NUMBER: PB248P1

TELECOMMUNICATION INFORMATION:

TELEPHONE: (240) 314-1224
 TELEFAX: (301) 309-8439

INFORMATION FOR SEQ ID NO: 725:

SEQUENCE CHARACTERISTICS:
 LENGTH: 1440 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 725:

US-08-956-171E-725

Query Match 65.1%; Score 1432; DB 4; Length 1440;

Best Local Similarity 99.7%; Pred. No. 2e-287;

Matches 1432; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 29 AAATACAAAATTTAAATATATTACCTTATTAGAAAAGTCGTATATATGAGGTGTACAAA 88

Db 4 AAATNCANTATTNAAATATATTACCTTATTAGAAAAGTCGTATATGAGGTGTACAAA 63

Qy 89 TGACGCAAAATTTTAAATAGTAGAAGTAGAAGTGAACAACTTAGCAAGATTTCTTGAATTGGAAC 148

Db 64 TGACGCAAAATTTTAAATAGTAGAAGTAGAAGTGAACAACTTAGCAAGATTTCTTGAATTGGAAC 123

Qy 149 TCACACATGAAAATTTACAAATGTGGACACAGAGTATGATGGACAGACGGTTTAGATAAAG 208

Db 124 TCACACATGAAAATTTACAAATGTGGACACAGAGTATGATGGACAGACGGTTTAGATAAAG 183

Qy 209 CGCTTAGCCATTACTATGATTTAATCATATTTAGATTTAAATGTTCCGCTCAATTAATGGCT 268
 Db 184 CGCTTAGCCATTACTATGATTTAATCATATTTAGATTTAAATGTTCCGCTCAATTAATGGCT 243
 Qy 269 TAGAAATTTGTGCAAAATTTAGACAAACAACTCTACACCTATCATTTAATTAACAGGGA 328
 Db 244 TAGAAATTTGTGCAAAATTTAGACAAACAACTCTACACCTATCATTTAATTAACAGGGA 303
 Qy 329 AAAGTGATAGTATGACAAAGTTGCTGGGCTTGATTAACGGTGACAGATTTATATAGTTA 388
 Db 304 AAAGTGATAGTATGACAAAGTTGCTGGGCTTGATTAACGGTGACAGATTTATATAGTTA 363
 Qy 389 AGCGGTTTGATTTGAAAGAACTTTTAGCAAGAAATTCGTGCAATTTTACGTCGTGAGCCAC 448
 Db 364 AGCGGTTTGATTTGAAAGAACTTTTAGCAAGAAATTCGTGCAATTTTACGTCGTGAGCCAC 423
 Qy 449 AAAGGATATATCGATGTCACCGGTTTCAATTTGATTAAGAACGCTTTTAAAGTACGG 508
 Db 424 AAAGGATATATCGATGTCACCGGTTTCAATTTGATTAAGAACGCTTTTAAAGTACGG 483
 Qy 509 TAAATGCGCAGAAATTTGAATTTAACAACAAACAGAGTATGATTTACTATATCTTCTAGCTG 568
 Db 484 TAAATGCGCAGAAATTTGAATTTAACAACAAACAGAGTATGATTTACTATATCTTCTAGCTG 543
 Qy 569 AAAATAAAACCATTGTTATGCAACGGGAACAAATTTTAAATCATGTATGGGTTTATAATA 628
 Db 544 AAAATAAAACCATTGTTATGCAACGGGAACAAATTTTAAATCATGTATGGGTTTATAATA 603
 Qy 629 GTGAAGTAGAAAACAAATGTCGTAGATGTTTATATAAGATATTTACGAAAACAAAGTTAAAC 688
 Db 604 GTGAAGTAGAAAACAAATGTCGTAGATGTTTATATAAGRTATTTACGAAAACAAAGTTAAAC 663
 Qy 689 CATACGATCGTGACAAATGATTTGAAACAGTTCTGCGGCTTGGGTATGATGATGATGAC 748
 Db 664 CATACGATCGTGACAAATGATTTGAAACAGTTCTGCGGCTTGGGTATGATGATGATGAC 723
 Qy 749 AAAACGTAATTTGCGCAATTAATCTGATTTTGTACCAACGATGATTAACGTTGTGACGAT 808
 Db 724 AAAACGTAATTTGCGCAATTAATCTGATTTTGTACCAACGATGATTAACGTTGTGACGAT 783
 Qy 809 ATTTTGTGTTTGTAAATTTATTTTCTTGAAAGATACACTGCAATATAGTAGCT 868
 Db 784 ATTTTGTGTTTGTAAATTTATTTTCTTGAAAGATACACTGCAATATAGTAGCT 843
 Qy 869 TGATGATGACGACGAGCTCAAGGATATTAATTAATTTTATTTTCAATCTAGGCTGTAA 928
 Db 844 TGATGATGACGACGAGCTCAAGGATATTAATTAATTTTATTTTCAATCTAGGCTGTAA 903
 Qy 929 AGATATATCTGCATTAGACTTTGAATGATCTTTAGGTAATTTTCAAGAGATAATTTTAA 988
 Db 904 AGATATATCTGCATTAGACTTTGAATGATCTTTAGGTAATTTTCAAGAGATAATTTTAA 963
 Qy 989 TGATGACATAAATAAATTTATTTGAGACATCGAATGATAACACAGTGAGGTGAAC 1048
 Db 964 TGATGACATAAATAAATTTATTTGAGACATCGAATGATAACACAGTGAGGTGAAC 1023
 Qy 1049 AGGTTATGAACACCGTTATTTTGACCGGTAATAAAGGCTTATAAGGCTTGAATA 1108
 Db 1024 AGGTTATGAACACCGTTATTTTGACCGGTAATAAAGGCTTATAAGGCTTGAATA 1083
 Qy 1109 TTTAATTTTAAAGAACCAATTTACACGCAAGATTTTCAAGGGGTATAGCTTGTAAATTTCA 1168
 Db 1084 TTTAATTTTAAAGAACCAATTTACACGCAAGATTTTCAAGGGGTATAGCTTGTAAATTTCA 1143
 Qy 1169 TTCACTAGAAAATTTATGATAACATCGTAAATCATTTGATATATCATTTGCGTGGCATTTGG 1228
 Db 1144 TTCACTAGAAAATTTATGATAACATCGTAAATCATTTGATATATCATTTGCGTGGCATTTGG 1203
 Qy 1229 AGTGATTTGCAACAAATTTAATTAAGTGAACATCGTAAATCATTTGATATATCATTTGCAACAAATTAATA 1288
 Db 1204 AGTGATTTGCAACAAATTTAATTAAGTGAACATCGTAAATCATTTGATATATCATTTGCAACAAATTAATA 1263

QY 1289 ACCGCTTGTCAGTTTATCAAAATTAATGATTGAGATTCGACGAGATGGTTTCAAATAA 1348
DB 1264 ACCGCTTGTCAGTTTATCAAAATTAATGATTGAGATTCGACGAGATGGTTTCAAATAA 1323
QY 1349 ATTGCAATTAATAACAAATTTATGAAGAAATAGATAATTTAGCAAAATACGTTTAAATGAGAT 1408
DB 1324 ATTGCAATTAATAACAAATTTATGAAGAAATAGATAATTTAGCAAAATACGTTTAAATGAGAT 1383
QY 1409 GATGAGCAATTAATAACAAATTTATGAAGAAATAGATAATTTAGCAAAATACGTTTAAATGAGAT 1465
DB 1384 GATGAGCAATTAATAACAAATTTATGAAGAAATAGATAATTTAGCAAAATACGTTTAAATGAGAT 1440

RESULT 3
US-08-771-110-1/c
; Sequence 1, Application US/08771110
; Patent No. 6084086
; GENERAL INFORMATION:
; APPLICANT: Wallis, Nicola G.
; APPLICANT: Hodgson, John E.
; TITLE OF INVENTION: No. 6084086el Compounds
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Smithline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-2799
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/771,110
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Gimmi, Edward R.
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: P31316
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 736 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
US-08-771-110-1

Query Match 33.4%; Score 736; DB 3; Length 736;
Best Local Similarity 100.0%; Pred. No. 1.5e-143; Indels 0; Gaps 0;
Matches 736; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 TAGAAAAATCAAAATTTAAATATATATACCTTATTAGAAAAAGTCGTAATATGAGGTGT 83
DB 736 TAGAAAAATCAAAATTTAAATATATATACCTTATTAGAAAAAGTCGTAATATGAGGTGT 677
QY 84 ACAATGACGCAATTTTAAATAGTAGAAGATGAACAAACTTAGCAAGATTTCTTGATTT 143
DB 676 ACAATGACGCAATTTTAAATAGTAGAAGATGAACAAACTTAGCAAGATTTCTTGATTT 617
QY 144 GGAATCACATCAATGAAATTTACATGTGGACACAGAGTATGATGCAAGACGTTTGA 203
DB 616 GGAATCACATCAATGAAATTTACATGTGGACACAGAGTATGATGCAAGACGTTTGA 557
QY 204 TAAAGCGCTTAGCCATTACTATGATTTAATCATATATAGATTTAATGTTGCCGTCATTA 263
DB 556 TAAAGCGCTTAGCCATTACTATGATTTAATCATATATAGATTTAATGTTGCCGTCATTA 497

QY 264 TGSCCTTAGAAAATTTGTGCGAAAATTAGACAAACAAATCTACACCTATCATTTAATATAC 323
DB 496 TGSCCTTAGAAAATTTGTGCGAAAATTAGACAAACAAATCTACACCTATCATTTAATATAC 437
QY 324 AGCGAAAAGTGATACGTTATGACAAAAGTTGCTGGGCTTGATTCGCGTGCAGACGATTATAT 383
DB 436 AGCGAAAAGTGATACGTTATGACAAAAGTTGCTGGGCTTGATTCGCGTGCAGACGATTATAT 377
QY 384 AGTTAAGCCGTTTGATTAATGAAGAACTTTAGCAAGAAATTCGTCGAATTTTACGTCGTCA 443
DB 376 AGTTAAGCCGTTTGATTAATGAAGAACTTTAGCAAGAAATTCGTCGAATTTTACGTCGTCA 317
QY 444 GCCACAAAAGGATATTATCGATGTCACCGTATTACAAATTTGATAAGAACGCTTTTAAAGT 503
DB 316 GCCACAAAAGGATATTATCGATGTCACCGTATTACAAATTTGATAAGAACGCTTTTAAAGT 257
QY 504 GACGTTAAATGGCGCAGAAAATTTGAATTAACAAAAACAGAGTATGATTTACTATATCTTCT 563
DB 256 GACGTTAAATGGCGCAGAAAATTTGAATTAACAAAAACAGAGTATGATTTACTATATCTTCT 197
QY 564 AGCTGAAAATAAAACCATGTTATGCAACGGGAAACAAATTTTAAATCATGTTATGGGTTA 623
DB 196 AGCTGAAAATAAAACCATGTTATGCAACGGGAAACAAATTTTAAATCATGTTATGGGTTA 137
QY 624 TAATAGTGAAGTAGAACAATAATGTCGTAGATGTTTATATAAGATATTTACGGAACCAAGTT 683
DB 136 TAATAGTGAAGTAGAACAATAATGTCGTAGATGTTTATATAAGATATTTACGGAACCAAGTT 77
QY 684 AAAACCATACGATCGTGACAAAATGATGAAACAGTTTCGTCGGCTTGGGTATGATGATACG 743
DB 76 AAAACCATACGATCGTGACAAAATGATGAAACAGTTTCGTCGGCTTGGGTATGATGATACG 17
QY 744 ATGACAAAACGTTAAAT 759
DB 16 ATGACAAAACGTTAAAT 1

RESULT 4
US-09-642-000-3/c
; Sequence 3, Application US/09642000
; Patent No. 6548281
; GENERAL INFORMATION:
; APPLICANT: Wallis, Nicola G.
; Shilling, Lisa K.
; Mooney, Jeffrey L.
; Debouck, Christine
; Zhong, Yiyi
; Jaworski, Deborah D.
; Wang, Min
; Throup, John P.
TITLE OF INVENTION: Histidine Kinase
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dechert, Price & Rhoads
STREET: 4000 Bell Atlantic Tower, 1717 Arch St
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103-2793
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/642,000
FILING DATE: 18-Aug-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/006,627
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:

NAME: Falk, Stephen T
REGISTRATION NUMBER: 36,795
REFERENCE/DOCKET NUMBER: GM10127
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-994-2488
TELEFAX: 215-994-2222
TELEX: <unknown>
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 736 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-642-000-3

Query Match 33.4%; Score 736; DB 4; Length 736;
Best Local Similarity 100.0%; Pred. No. 1.5e-143;
Matches 736; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 24 TAGAAAAATACAAAAATTTAAAAATATATTACCTTTATTAGAAAAAGTCGTAATATAGAGTGT 83
DB 736 TAGAAAAATACAAAAATTTAAAAATATATTACCTTTATTAGAAAAAGTCGTAATATAGAGTGT 677
QY 84 ACAATAGCGCAAAATTTTAATAGTAGAAGATGAACAAAACTTTAGCAAGATTTCTTTGAATT 143
DB 676 ACAATAGCGCAAAATTTTAATAGTAGAAGATGAACAAAACTTTAGCAAGATTTCTTTGAATT 617
QY 144 GGAATCTACACATGAAATTTACAAATGTGGACACAGAGTATGATGCAAGACGGTTTAGA 203
DB 616 GGAATCTACACATGAAATTTACAAATGTGGACACAGAGTATGATGCAAGACGGTTTAGA 557
QY 204 TAAAGCGCTTAGCCATTACTATGATTAAATCATATTAGATTAAATGTTGCCGTCAATTAA 263
DB 556 TAAAGCGCTTAGCCATTACTATGATTAAATCATATTAGATTAAATGTTGCCGTCAATTAA 497
QY 264 TGGCTTAGAAATTTGTCGCAAAATTTAGACAAACAACAACTACACCTATCATTAATTAAC 323
DB 496 TGGCTTAGAAATTTGTCGCAAAATTTAGACAAACAACAACTACACCTATCATTAATTAAC 437
QY 324 ACCGAAAAGTGTAGTATGACAAAGTTGCTGGGCTTGAATACGGTGCAGCATTATAT 383
DB 436 ACCGAAAAGTGTAGTATGACAAAGTTGCTGGGCTTGAATACGGTGCAGCATTATAT 377
QY 384 AGTTAAGCGCTTGTATTTGAAGAAGCTTTTAGCAAGAATTCGTGCAATTTTACGTCGCA 443
DB 376 AGTTAAGCGCTTGTATTTGAAGAAGCTTTTAGCAAGAATTCGTGCAATTTTACGTCGCA 317
QY 444 GCCACAAAAGGATATTATCGATGTCACACGGTATTACAAATGTGATAAGAAGCTTTTAAAGT 503
DB 316 GCCACAAAAGGATATTATCGATGTCACACGGTATTACAAATGTGATAAGAAGCTTTTAAAGT 257
QY 504 GACGTTAAATGCGCAGAAATTTGAATTAACAAAAACAGAGTATGATTTACTATATCTTCT 563
DB 256 GACGTTAAATGCGCAGAAATTTGAATTAACAAAAACAGAGTATGATTTACTATATCTTCT 197
QY 564 ACCTGAAATTAATAAACCTGTTATGCAAGCGGAACAAATTTTAAATCATGATGCGGTTA 623
DB 196 ACCTGAAATTAATAAACCTGTTATGCAAGCGGAACAAATTTTAAATCATGATGCGGTTA 137
QY 624 TAATAGTGAAGTAGAACAATGTCGTAGATGTTTATATAGATATTTTACGAAACAAGTT 683
DB 136 TAATAGTGAAGTAGAACAATGTCGTAGATGTTTATATAGATATTTTACGAAACAAGTT 77
QY 684 AAAACCATACGATCGTGACAAATGATTTGAAACAGTTTCGTGCGGTTGGGTATGATACG 743
DB 76 AAAACCATACGATCGTGACAAATGATTTGAAACAGTTTCGTGCGGTTGGGTATGATACG 17
QY 744 ATGACAAAACGTAAT 759
DB 16 ATGACAAAACGTAAT 1

RESULT 5
US-09-134-001C-313
; Sequence 313, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 313
; LENGTH: 1416
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-313
Query Match 33.3%; Score 732.8; DB 4; Length 1416;
Best Local Similarity 70.7%; Pred. No. 7.7e-143;
Matches 990; Conservative 0; Mismatches 407; Indels 3; Gaps 1;
QY 700 GACAAAATGATTGAAACAGTTTCGGGCGTTGGGTATGTGATACGATGACAAAACGT---A 756
DB 2 GAAAATCCATAGAAACAGTACGTGGCGTAGGGTATGTGATTCGATGATGTAAGCGCCAAA 61
QY 757 AATTGCGCAATPACTGGATTATTGTTACCCAGTATGATACGTTTGTCCGATATTTTGT 816
DB 62 AATTAATAATAATAANTGGATGCTTATTACGACGCTCATTACCTTCACGACAATATTACITT 121
QY 817 TTTGTTAAATATTATTTTCTTTGAAAGATACACATGATGATGAGCTTGATGATG 876
DB 122 TCTGCTTAATTTATTTTCTTTTAAAGATACCTTTACGAAAGTAGTGAATTTGACGAAG 181
QY 877 CAGAACGAAGCTCAAGCGATTAATAATTTTATTTTCAAGCCTGTTAAAGATATAT 936
DB 182 CTGAAGAAGTTCAATGATATATCCCAATCTGTTCCATCTTAAATCTTTAAGTGATATAT 241
QY 937 CTGCATTAGACTTGAATGCAATCTTTTAGGTAAATTTTCAAGAGATATATTTATGATGAGC 996
DB 242 CTGCATTGGATTTAAATGCAATCTTTAGAAAATTTTCAAGAAATATTGATTTATGATGATA 301
QY 997 ATAAATAATAATTTTGAAGCATCGAATGATATACAGTGAGAGTTGAACACGATTATG 1056
DB 302 AGGTTAGGAAGTTAAATCAAAACATCAAAATGATAATACACTTGTCTTATGATAACAAAATTG 361
QY 1057 AACACCGTTATTTTACCGGCTAATAAAAAACGCTATAAAGGCATTGAAATTTAAATTA 1116
DB 362 ATTTCAAAACATCTGAAAGTATATATATTCAAAGAGCCACGGTATTAATTTACTTAGTAA 421
QY 1117 TTAAGAAGCAACCAATTTACAAACGCAAGATTTCAAGGGTATAGCTTGTAAATTCATTCACTAG 1176
DB 422 TAACTGAACCTTACGTTTCGAAAGATTTTCTGGATACAGTGATATTAGTCCATCTCTTC 481
QY 1177 AAAATTATGATAACATCGTAAAATCATTTGATATATCATTTGCGCTGGCATTTTGGAGTGATTG 1236
DB 482 AAAATTATGATAATCTCGTTAAATCACTTTATATAGTTGCACTTGTCTTTTGGATTAATTG 541
QY 1237 CAACAAATTAACCTGCCACAAATCAGTTATGTTTCAACACAAATTTACTTAAACCGCTTG 1296
DB 542 CAACCAATTTTACTGCGCGGTGAGTTATATCTTTTCTTCAAAATTTACTTAAACCGTAG 601
QY 1297 TCAGTTTATCAAAATAAATGATTGAGATTCGACGAGATGTTTTCACAAATAAATTTGCAAT 1356
DB 602 TTACAAATGTCCAATAAATGAATCAAAATGAGAGATGTTTTCACAAATAAATTTGCAAT 661
QY 1357 TAAATACAAATTTAGAGAAATAGATTAATTTAGCAAAATACGTTTAAAGATGATGAGCC 1416
DB 662 TAACTACAAATTTAGAGAAACAGATAATTTAAATTTGATCTTTTAAATGAAATGATGATC 721

```
Qy 1417 AAATTGAAGAAATCATTTAATCAACAAAGACAAATTTGTTGAAGATGCGTCACATGAATTAC 1476
Db |||||
Qy 722 AAATAGAAGATCTTTAATCAGCAACGTCATTTGTCGAGGATGCTTCACAGAAATTA 781
Db |||||
Qy 1477 GAACACCAATTAACAATTAATCAAGGTCAATTTAAATTTGAATTCAGCGATGGGGAAAAAAG 1536
Db |||||
Qy 782 GAAGCCCACTGCAGATTAATCAAGGTCAATTTAAATTTAATCCAACGTTGGGGAAAAAAG 841
Db |||||
Qy 1537 ACCGAGCAGTATTAGAAGAAATCGTTAAATATTTCTATTGAAGAAATGAATGATCATATA 1596
Db |||||
Qy 842 ATCCAGCAGTTTGAAGAAATCTTTGAATATTTCAATTTGAAGAAATGAATGATCATATA 901
Db |||||
Qy 1597 AATTAGTCGAAGAAATTTACTTGAATTTGACTAAGAGAGATGTAATGCAATTTCTTCTGAAG 1656
Db |||||
Qy 902 AACTTGTGGAAGAAATCTTTTACTTACCAAGATAGAGTCAATCAATGTTTGGAT 961
Db |||||
Qy 1657 CGCAGACGTCATTAATGATGAATTCGCTCGCGAATACATCTAATTAATAAATGCAATTCG 1716
Db |||||
Qy 962 GTGAAATGTAGACGTAATAGCAGATTTCAATCACGTTGAGTCACTGCAACACCTAC 1021
Db |||||
Qy 1717 ATCTGATTAATCAATTTGATACGATCTGACATCTAATAAATCTAGAAATTAATAATGAAC 1776
Db |||||
Qy 1022 ATCCAGATTAATCTTTTGAACACATCTTGCTACTAAGCCCTATCCAATTAATAATTAACC 1081
Db |||||
Qy 1777 CTCAATCAATTCGAACAAATTTATTTTAAATCTTTTATTTGATAATGCAATCAATATGATGTA 1836
Db |||||
Qy 1082 GTCAATCAGTTTGAACAACTCTTACTCATATTTATTTGATGATGCAATGATGATGACACTG 1141
Db |||||
Qy 1837 AGAATAGAAATTAAGTTGAAGCAAGGTTTAAATAAATAAGCAAAATAATTAATGAATTA 1896
Db |||||
Qy 1142 AACATAAGCAATTAATAATTTTACTCAACTTAAATAAATAAATGATGATGATGATTA 1201
Db |||||
Qy 1897 CAGATCATGGAATTTGATTTCCAGGAGGATCAAGATTTCAATTTTTCGCTGCTTTTATC 1956
Db |||||
Qy 1202 CTGATCATGATGGGTATACCAAGCTGACTTAGAATTTATCTTTGATAGATTTTATC 1261
Db |||||
Qy 1957 GAGTGGATAATCTCGTTCAAGAAAGTCAAGCGGTAATGGAATGCTGATTTCTATTGCTC 2016
Db |||||
Qy 1262 GTGTAGATAATCAGTGCTCGTAGTCAAGAGGCAATGGATAGGATCAATGAGCGG 1321
Db |||||
Qy 2017 AAAAAATCAATTAATTAACGGAGGATCGATTAATAATTAAGTGAATTAATAAAGGAA 2076
Db |||||
Qy 1322 AAAAAATTTGCAACTTAAACGGTGTATGATTCAGGTAGAAAGTGAACACTACAAAGTACA 1381
Db |||||
Qy 2077 CAACGTTTAAATCATATTT 2096
Db |||||
Qy 1382 CGACTTTCAAAATCAGTTTT 1401
Db |||||
```

RESULT 6

US-08-956-171E-379
; Sequence 379, Application US/08956171E
; Patent No. 6593114

GENERAL INFORMATION:

APPLICANT: Charles Kunsch

Gil H. Choi

Patrick S. Dillon

Craig A. Rosen

Steven C. Barash

Michael R. Fannon

TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences

NUMBER OF SEQUENCES: 5256

CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue

CITY: Rockville

STATE: Maryland

COUNTRY: USA

ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage

COMPUTER: HP Vectra 486/33

OPERATING SYSTEM: MSDOS version 6.2

SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/956,171E

FILING DATE: 20-Oct-1997

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/009,861

FILING DATE: January 5, 1996

APPLICATION NUMBER: 08/781,986

FILING DATE: January 3, 1997

ATTORNEY/AGENT INFORMATION:

NAME: Mark J. Hyman

REGISTRATION NUMBER: 46,789

REFERENCE/DOCKET NUMBER: PB248P1

TELECOMMUNICATION INFORMATION:

TELEPHONE: (240) 314-1224

TELEFAX: (301) 309-8439

INFORMATION FOR SEQ ID NO: 379:

SEQUENCE CHARACTERISTICS:

LENGTH: 4858 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 379:

US-08-956-171E-379

Query Match 30.6%; Score 673.2; DB 4; Length 4858;

Best Local Similarity 99.6%; Pred. No. 2.1e-130;

Matches 675; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1524 TGGGGAAAAAAGACCCAGCAGTATTAGAGAAATCGTTAAATATTTCTATTGAAGAAATG 1583

Db 1 TGGGGAAAAAAGACCCAGCAGTATTAGAGAAATCGTTAAATATTTCTATTGAAGAAATG 60

Qy 1584 AATCGTATCATATAAATTTAGTCGAAGAAATTAATTTGAATGCTAAAGGAGATGTAATGAC 1643

Db 61 AATCGTATCATATAAATTTAGTCGAAGAAATTAATTTGAATGCTAAAGGAGATGTAATGAC 120

Qy 1644 ATTTCTTCGAGCGGAGACCGTGCATATTAATGATGAATTCGTCGCGAATACACTCA 1703

Db 121 ATTTCTTCGAGCGGAGACCGTGCATATTAATGATGAATTCGTCGCGAATACACTCA 180

Qy 1704 TTAACAATTTGATCTGATTAATCAATTTGATACGATCTGACATCTAAAAATCTAGAA 1763

Db 181 TTAACAATTTGATCTGATTAATCAATTTGATACGATCTGACATCTAAAAATCTAGAA 240

Qy 1764 ATTAATAAGAACTCATCAATTCGAACAAATTTATTTTAACTTTTATTTGATGCAATC 1823

Db 241 ATTAATAAGAACTCATCAATTCGAACAAATTTATTTTAACTTTTATTTGATGCAATC 300

Qy 1824 AATATGATGTGAAGAAATTAAGAAATTAAGTTAAGCAAGGTTAAAAATTAAGCAAAA 1883

Db 301 AATATGATGTGAAGAAATTAAGAAATTAAGTTAAGCAAGGTTAAAAATTAAGCAAAA 360

Qy 1884 ATAAATGAATTTACAGATCATGAATTTGGTATTTCAGAGGAAGATCAAGATTTTCAATTTT 1943

Db 361 ATAAATGAATTTACAGATCATGAATTTGGTATTTCAGAGGAAGATCAAGATTTTCAATTTT 420

Qy 1944 GATCGCTTTTATCGAGTGGATAAATCTCGTTCAAGAAAGTCAAGCGGTAATGAGCTCGGA 2003

Db 421 GATCGCTTTTATCGAGTGGATAAATCTCGTTCAAGAAAGTCAAGCGGTAATGAGCTCGGA 480

Qy 2004 TTATCTATTGCTCAAAAATCATTTCAATTAACGGAGGATCGGATTAATAATTAAGTGAA 2063

Db 481 TTATCTATTGCTCAAAAATCATTTCAATTAACGGAGGATCGGATTAATAATTAAGTGAA 540

Qy 2064 ATTAATAAGGAACAAACGTTTAAAAATCATATTTTAACTGCTGAGACGCTCAATCAAG 2123

Db 541 ATTAATAAGGAACAAACGTTTAAAAATCATATTTTAACTGCTGAGACGCTCAATCAAG 600

Qy 2124 TCATAGATCAATTTTAAAGTACACATTTAGCTGTGCTAATGATTAAGAACAACTATAA 2183

Db |||||

Db 601 TCATAGGATCAATTTTAAAGTACACATTAGCTGTGACTAATGTATATAAGAACAACTATAA 660
Qy 2184 AACAAATAAACAGTGGTT 2201
Db 661 AACAAATAAACAGTGGTT 678
RESULT 7
US-09-134-001C-375
; Sequence 375, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 375
; LENGTH: 669
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-375
Query Match 18.3%; Score 401.8; DB 4; Length 669;
Best Local Similarity 75.0%; Pred. No. 1.8e-74;
Matches 502; Conservative 0; Mismatches 167; Indels 0; Gaps 0;
Qy 79 GGTGTACAAATCAGCGAAATTTTAAATAGTAGAAGATGAACAAACTTAGCAAGATTCTT 138
Db 1 GGTGCAATAGCAAAATATTTTAAATTTAGTAAGATGACAAATCTTCTAGATTATA 60
Qy 139 GAATTGGAACACATGCAATTTTAAATGTGGACACAGAGTATGATGGCAAGACGGT 198
Db 61 GAGCTTGAGTTAACTCATGAAATTTACTGTTGATATAGAGATGATGGAAGTGGGT 120
Qy 199 TTAGATAAGCGCTTAGCCATTACTATGATTTAATCATATTTAGATTATTTGTCGGTCA 258
Db 121 TTGATAAAGCAATTTCTAAGCCCTATGATTTTATATATATATTAGACTTAATGCTTCCAAAC 180
Qy 259 ATTAATGCTTAGAATTTGTCGCAAAATTTAGACAACTATGACAACTATCACCCTATCATATA 318
Db 181 ATAAATGCTTAGAATTTTGTAGCAAAATTTCTGCAAAACAACTACTCCCAATATATATC 240
Qy 319 ATTACAGCGAAAGTGATACGATGATGACAAAGTTGCTGGGCTTGATTACGGTGCACAGAT 378
Db 241 ATTACTGCGAAAGCGAGACATATGATAAGTAGCTGGATTGGACTATGGGCGAGATGAC 300
Qy 379 TATATAGTTAGCCGTTTGATTTAGAGAACTTTTAGCAAGAACTTTGCGCAATTTTACGT 438
Db 301 TACATTGTAAACCCCTTTGATATAGAAGAAATGCTTGCAAGAAATGAGCGGATTTGGCG 360
Qy 439 CGTCAGCGCAAAAGGATTTATCGATGTCACGGTATTACAACTGATGAAGACGCTTTT 498
Db 361 AGACAGCGAGATAAAGATGTTTATGATATCAATGATATCAATGATGATGAAGATGCCCTTT 420
Qy 499 AAAGTCAGCGTAAATGGCGCGAGAAATTTGAATTAACAAAAACAGAGTATGATTTACTATAT 558
Db 421 AAAGTTACTGTTAATGGCCATCAATTAGAATTAACATAACAGATACGATTTACTATAT 480
Qy 559 CTCTAGCTGAAATAAATAACATGTTATGCAACGGGAACAAATTTTAAATCATGATGCG 618
Db 481 GTTTTAGCTGAAATCGTAACCAACGCTCATGCAACGTCGAACAAATTTCTCGATCAGCTATGG 540
Qy 619 GGTATATAGTGAAGTGAACAAATGCTCGTAGATGTTTATATAGATATTTTACGAAC 678
Db 541 GGGTATAATAGTGAAGTGAACAAATGCTCGTTGATGTTTATATTCGTTATTTAGCTAAT 600

Qy 679 AAGTTAAACCATACGATCGTGACAAAATGATGAACACAGTTCTGCGGTTGGGTATGTG 738
Db 601 AAACCTCAACCTTTTAAATAAAGAAAATCCATAGAAACAGTACCTGCGTAGGTATGTG 660
Qy 739 ATACGATCA 747
Db 661 ATTCGATGA 669
RESULT 8
US-09-107-532A-2777
; Sequence 2777, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 2777:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1566 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME:
; LOCATION: (B) LOCATION 1...1566
; SEQUENCE DESCRIPTION: SEQ ID NO: 2777:
US-09-107-532A-2777
Query Match 9.4%; Score 206.2; DB 4; Length 1566;
Best Local Similarity 51.4%; Pred. No. 6.2e-34;
Matches 528; Conservative 0; Mismatches 493; Indels 6; Gaps 2;
Qy 1076 CGTAATAAAAAACGCTATAAGGCATTGAATATTTAATTATTAAAGAACCAATTACAAC 1135
Db 513 CGTTACTAGAGAGAAAACCTGCTTTTATCTATCCAGCTATATATTTCTAAGGTAC 572
Qy 1136 GCAAGATTTCAAGGGTATAGCTTTGTTAAATTCATTCATAGAAAATATGATAACATCGT 1195
Db 573 AAGAGAAAAGATCGGCTATGCGCAAGCTTCTATGAATTTATCTTCTTCTATGATATTCG 632

QY 1196 AAAATCATTTGATATCATTTGGCGTGGCAATTTGGAGTGAATTCACAAATTAATTAACCTGCCAC 1255
DB 633 AAATAAGCTATTAATGACTCTTAATGTTTGAAGAGTTGTTCTTTGATCTTAAGTAGTAT 692
QY 1256 AATCAGTTATGATTTTCAACACAAATTAATTAACCGCTTCTGAGTTTCAATTAATAAAT 1315
DB 693 TTTAGTTCTTTTATCTTTATCTTTTCTTTAAACCACTCAAGTACTCAGAGATACAT 752
QY 1316 GATTGAGATTCGACGAGATGTTTTCAAAATA---AATTGCAATTAATAACAAATTAATGA 1372
DB 753 GGATCTATACGTAAGATCCGCAATCAGATATCCACATGCTCAATTCGACAAATGA 812
QY 1373 AGAATAGATTAATTAGCAATAACGTTTAATGAGATGATGAGCCAAATTAAGAAATCAT 1432
DB 813 TGAATGGCAGATTTAGCTGAATATTAATGAGATGCTTGCACGATGCTTGTATAT 872
QY 1433 TAATCAACAAAGACAAATTTGTTGAAGATGCTGCATCAATTAATGACACCAATTAACAAAT 1492
DB 873 TGAACAAAGAGCAATTCGTTGAAGACGATATCTCATGAGTTAAGGACACCTGTTGCAAT 932
QY 1493 TATTCAAGGTCAATTAATTTGATTGACGATGGGAAACCAACAGCAGCATTTAGA 1552
DB 933 TATTGAAGGATCTTAGTCTGCTAAATCGTTGGGGAAGATGATCTGAGATTTAGA 992
QY 1553 AGAATCGTTAAATATTTCTATTGAAGAAATGAATCGTATCATAAATTAATGTCGAGAAAT 1612
DB 993 AGAATCATTAAGCCAGCATGCAAGAAATCAGCCGATGAAAGCTTAGTTCAAGAT 1052
QY 1613 ACTTGAATGACTAAAGGAGATTAATGACATTTCTTCAAGCCGAGACCGTGCATAT 1672
DB 1053 GCTAGATTAATCAAGAGACAAAGTAGATGTTTCATTTATGCGAATGAACCAACTAAGG 1112
QY 1673 TAATGATGAATTCGCTCGCAATACACTCAATTAACAAATTTGATCTGATTAATCAAT 1732
DB 1113 AAAGAGTCACTTCAAGTTTATATTAATTTCAAGATTAATTTCAAGATTTGAT 1172
QY 1733 TGATACGATCTGACATCAAAATCTAGAAA---TTAAATGAACCTCATCAATTCGA 1789
DB 1173 CACTTTAGATGATGATTTGACAAAGAGTGAACCTGCAAAATTTATCGTAATCACTTTGA 1232
QY 1790 ACAATTTTATTTTATTTTATGATTAATGCAATCAATATGATGTAAGATTAAGAAAT 1849
DB 1233 ACAATTAATCAATATCATATTGGATAACGAGTAAGATTTCAACTACACGAAAGAGT 1292
QY 1850 TAAAGTTAAGACAGGTTAAATAATAAGCAAAATAATTAATTAACATCATGGAAT 1909
DB 1293 CCATATCTTATTTCTTCCACTCTGATGAATTTGAGATCGCGATTCAGACTTTGGCA 1352
QY 1910 TGGTATTCAGAGGAAGATCAAGATTTTCAATTTTGTGATCGTTTATCGAGTGAATAATC 1969
DB 1353 AGGAATACCAAGAGACCTTAATAAGATCTTCAATCGTTTATCGAGTGAATAAGC 1412
QY 1970 TCGTTCAAGAGTCAAGCGGTANTGACTCGGATTTATCTTATGCTCAAAATCATTTCA 2029
DB 1413 AAGAGTAGAACGAGGTTGGAATGCTCTGGGTTGCTATCGCTAAACATCGTTGA 1472
QY 2030 ATTAACCGGAGTTCGATTAATAATTAAGTGAATTAATAAGGAACCAACGTTTAAAT 2089
DB 1473 AACTATAAGGAGTATCTTAGCAGAAAGTGTAGTAGGCAAGGCAATTTTATAGAT 1532
QY 2090 CATATTT 2096
DB 1533 CTTTAT 1539

RESULT 9

US-09-134-000C-557
; Sequence 557, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO

; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 557
; LENGTH: 1128
; TYPE: DNA
; ORGANISM: Enterococcus faecalis
; US-09-134-000C-557

Query Match 9.4%; Score 205.8; DB 4; Length 1128;
Best Local Similarity 52.1%; Pred. No. 6.9e-34;
Matches 509; Conservative 0; Mismatches 462; Indels 6; Gaps 2;
QY 1126 CAATTAACAGCGAAGATTTCAAGGGTATAGCTTTTAATTCATTCACTAGAAAAATTTATG 1185
DB 113 CAAAAGAAAACACGAGAAAAGATTGGCTATATTCAAGCGTTTTATGAACCTTTCTCTTTT 172
QY 1186 ATAACATCGTAAATCATTTGATATCATTTCCGCTGGCATTTGAGTGAATTCGACAAATTA 1245
DB 173 ATGAAATTCGTAATCATTTTATTAACACTTTCTAGTTTTAGAGTGAATTTCTCTGATTG 232
QY 1246 TAACTGCCCAATCAGTTATGATTTTCAACACAAATTAATAACCGCTTGTGAGTTAT 1305
DB 233 TAAGTAGTCTTAGCTTTATCTCTCTTTCTTTTCTTTAAACCATTAAGAGTATGA 292
QY 1306 CAAATTAATGATTTGAGATTCGACGAGATGGTTTTCAAAAT---AAATTCGAATTAATA 1362
DB 293 GAGACAGATGGATACAAATCCGCAAGAGCCGCAATCAGATGCCACATGCGGAGATTA 352
QY 1363 CAAATTAAGAAATAGATTAATTTAGCAATAGCTTTTAATGAGATGATGAGCAATTTG 1422
DB 353 ATACGAGATGAGTTAGCAGATATCTCGAAATCTTTAATGAATTTGTAGATCGTATGA 412
QY 1423 AAGAATCATTTAATCAACAAAGACAAATTTGTTGAAGATCGCTCACATGAATTAACGAAC 1482
DB 413 GACGCTATATCGAACAAACGAGCAGTTTGTGAGATGTTTCCCATGAATTAAGAACGC 472
QY 1483 CATTACAAATTAATCAAGTCAATTTAAATTTGATTCAGCGATGGGAAAAAGAACCCAG 1542
DB 473 CCGTTCGATTTATGGAAGCCATTTAAACCTTTTAAATCGTTGGGGAAGACGATCCTG 532
QY 1543 CAGTATTAGAGATCGTTAAATATTTCTATTGAAGAAATGAATTCGTATCATTAATTTAG 1602
DB 533 AAATTTAGATGAATCATTTAAAGGCCAGTTTACAGAAATTTAGTCGTATGAAGATTTGG 592
QY 1603 TCGAAGAAATTAATTTGATTAATGATTAATGATTAATGATTAATTTCTTCTGAAGCGCAGA 1662
DB 593 TCCAGAAATGCTTGACCTTTACGCGCTGAAACAGTGGACACCCCAATTAATGAATGAAC 652
QY 1663 CGTGCAATTAATGATGAATTCGCTCGCAATACACTCATTTAAACCAATTTGATCCTCTG 1722
DB 653 GAACAGATGCTAAACAAAGTAGTCTACCAAGTATTTAAATTAATTTCAATTTGTTTATCCTG 712
QY 1723 ATTATCAATTTGATAGGATCTGACAT---CTAAAAATCTAGAAATTAATAATGAACCTC 1779
DB 713 AGTTTCATATCATATTAGATGACGATTTTACCAACCGAATTTAGCTGAAATCTATCGTA 772
QY 1780 ATCAATTCGAACAAATTAATTTTAAATCTTTTATGATAATGCAATCAATATGATGTGAAGA 1839
DB 773 ATCACTTTGAAATTTGCTAATTTATTTCTATTAGATTAATGCGATTAATTAATTAACCGATC 832
QY 1840 ATAAGAAAAATTAAGTTAAGACAAAGGTTAAAAAATAAGCAAAAAATTAATTTGAATTAACAG 1899
DB 833 GAAAAGAGTGATATTTCCATTTCTCGAACGATGAATGAATTTTGAATAGCTGTGCAAG 892
QY 1900 ATCATGGAATTTGATTTCCAGAGGAGATCAAGATTTTCAATTTTGTATCGCTTTTATCGAG 1959
DB 893 ATTTTGGTGAAGGAATCACTGAAGAAGACTTGGAGAAAAATATTTGATCGATTTCTATCGAG 952


```

Db 361 CAAAAGGAGATAGCAATATTATTACTATTGCTGATCTTGAACCTTGTAACTCTTCGTAAACGC 420
Qy 499 AAGTGCAGCGTAAATGGCGCAGAAATGAAATTAACAAAAACAGAGATGATGATTACTATAT 558
Db 421 CGCGTCACACGCTGGCAAGAATTGACCTACAGCTTAAGAGTTCGCAATGATGGA 480
Qy 559 CTTCTAGCTGAAATAAACAACATGTTATGCAACGGGAACAAATTTTAAATCATGTATGG 618
Db 481 TTATTATCGCTAGACGCTGGAGAGATCTCGCTGCTCTCAATTCCTCAAAATATGG 540
Qy 619 GGTATATATAGTGAATGAGTGAACAAATGCTGATGATGTTATATAGATATTTACGAAAC 678
Db 541 GATATGAATTTTATGATGATGATCAACAGTGTGAGAAAGTCGCTATTAACGCTCAAGAAAT 600
Qy 679 AAGTTTAAACCATGATCGTGACAAAAATGATTGAAACAGTTCGTGGCGTGGGTATGTCG 738
Db 601 AAAATAGATAGTAATTTTATGCTTAACTTAAGTTGATTCAAAACATCCGCGGAATGGGTATGTC 660
Qy 739 ATACGATGACAAAACG 754
Db 661 TTAGAGGTAGAAGATG 676

```

RESULT 14

```

US-08-956-171E-822
; Sequence 822, Application US/08956171E
; Patent No. 6593114
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; Gil H. Choi
; Patrick S. Dillon
; Craig A. Rosen
; Steven C. Barash
; Michael R. Fannon
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5256
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/956,171E
; FILING DATE: 20-Oct-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/009,861
; FILING DATE: January 5, 1996
; APPLICATION NUMBER: 08/781,986
; FILING DATE: January 3, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mark J. Hymen
; REGISTRATION NUMBER: 46,789
; REFERENCE/POCKET NUMBER: PB248P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (240) 314-1224
; TELEFAX: (301) 309-8439
; INFORMATION FOR SEQ ID NO: 822:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 760 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 822:
US-08-956-171E-822

```

```

Query Match 5.9%; Score 130.2; DB 4; Length 760;
Best Local Similarity 52.6%; Pred. No. 2.7e-18;
Matches 358; Conservative 0; Mismatches 288; Indels 35; Gaps 2;
Qy 90 GACCCAAATTTTANTAGTAGAAGATGAACAAAACCTTAGCAAGATTTCTTGAATTTGGAAC 149
Db 55 GAACGAAATCTTATCGTAGATGATGAGATAGAAATCAGAAGATTACTTAAATATGTATTT 114
Qy 150 CACACATGAAATATACAAATGTGGACACAGAGATGATGGCAACAGCGTTTGTAGATAAAGC 209
Db 115 AGAAGAGATCTTTTGAATCCATGAAGCAAGTAATGGCCAAAGAGCTTATGAACCTGC 174
Qy 210 GCTTAGCCATCTACTATGATTTAATCATATTAGATTTTAAATGTTCGGTCAATTAATAGCTT 269
Db 175 AATGGAATAAATATGCTTGCATCTACTAGATTTAATGTTCGCTGAAATGGATGGTAT 234
Qy 270 AGAATTTTTCGCAAAATTAGACAAACAATCTACACCTATCATTTAATATTACAGGAA 329
Db 235 CCAGTGGCAACTAAATTCGTGAACATAAACAACACCGATTTATTTGACTGCTAA 294
Qy 330 AAGTATACGTTATGACAAAGTTGCTGGCTTTGATTACGGTGCAGACGATTTATAGTTAA 389
Db 295 AGTGAAGAAACAACCGTGTGAGGTTTGAATCTGTTGCAGATGATTTATATCGTCAA 354
Qy 390 GCCGTTTATTTGAAGAAGCTTTTAGCAAGAAATTCGTGCAATTTTACGTCTCAGCCA-- 447
Db 355 ACCATTTTCCACCAAGAGAAGTAGTCTTAAAGAGTTTAAAGCACTTCTAAGAAGAACGCAATC 414
Qy 448 -----CAAAGGATATTATCGATGTCACGGTATTAC 479
Db 415 TACAACCTGTAGAACAAAGCGAACCTCACGACGATGATGATGAATTTTAAACATTTTAGA 474
Qy 480 AATTGATAAGAACCTTTTAAAGTGACGGTAAATGGCGCAGAAATTTGAATTTAAACAAAAC 539
Db 475 AATAGATTAATGATGCATCGCTACTTGTGATTAATCAAGAAGTTAATTTGACTCTCAA 534
Qy 540 AGAGTATGATTTACTATATCTTCTAGCTGAAAAATAAAAAACCATGTTATGCAACGGGAACA 599
Db 535 AGAGTACGAATTTAATATATTTAGCTAAACACCAAAATAAAGTATTTTGACCGTGAACA 594
Qy 600 AATTTAAATCATGTATGGGGTTATATAGTAGAAGTAGAACAATAATCGTAGATGTTTA 659
Db 595 ATTATTAAGAAGCTTTGGCAATTATGAATTTCTATGTTGATTTAAGAACAGTTGATACTCA 654
Qy 660 TATAAGATATTACGAACAAGTTAA-----AACCATACGATCGTGACAAAATGATTGAA 714
Db 655 TGTTAAACGCTTAGAGAAAAGTTAAATCGTGTCTAGTGAAGCTGCGCATATGATTCAA 714
Qy 715 ACAGTTCGTGGCGTTGGGTAT 735
Db 715 ACAGTCTGGGCGTTGGGTAT 735

```

RESULT 15

```

US-09-134-000C-2259
; Sequence 2259, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2259
; LENGTH: 696
; TYPE: DNA
; ORGANISM: Enterococcus faecalis

```

US-09-134-000C-2259

```
Query Match      5.9%; Score 129; DB 4; Length 696;
Best Local Similarity 54.0%; Pred. No. 4.7e-18;
Matches 301; Conservative 0; Mismatches 235; Indels 21; Gaps 1;

Qy 92 CGCAAAATTTTAATAGTAGAAGATGACAACTTTAGCAAGATTTCTTTGAAATTTGGAACCTCA 151
Db 11 CACATGTATTATTAATTTGAAGATGATCCAGCATTTGCTGATTTGCCAAAAGATTTATTAG 70

Qy 152 CACATGAAAATTACAATGTGGACACAGAGTATGATGGACAGAGCGTTTAGATAAAGCGC 211
Db 71 AAATTAACGATATGACTGTGACAAATTCACATGATGGAAAAAAGGATTTGGAAGCAGCTT 130

Qy 212 TTAGCCATTACTAGATTTAATCATATTAGATTTAATGTTGCCGTCAATTAATGGCTTAG 271
Db 131 TAAATGAGCCCTTTGATTTAATTTATTTTAGATGTGATGCTGCCAACTATGGATGGTTTG 190

Qy 272 AAATTTGTCGCAAAATTAGACAAACAATCTACACCTATCATTTAATTAACAGCGAAAA 331
Db 191 AAATCTGTCGTGCAATTCGAAAAAAGCAACGCCGCATCATGTTGTTCCGCTAAAA 250

Qy 332 GTGATACGTATGACAAAGTTGCTGGGCTTGATTAACGGTGCAGACGATTTATATAGTTAAGC 391
Db 251 AAGAAGATATTGTATAAAATTCGTGGGCTTGCTTAGGTGCGGATGATTATATAATCAAGC 310

Qy 392 CGTTTGATTTGAAGAATTTTAGCAAGATTCGTGCAATTTTACGTCGT----- 441
Db 311 CGTTTAGTCCCAATGAATTAGTCGCTCGAGCAAAAGCCCATATGAACCGTTATCAATGT 370

Qy 442 -----CAGCCACAAAAGGATATTATCGATGTCAACGGTATTACAAATTGATAAGA 490
Db 371 TAAGCAAGCAGACAAACCCACCAATTAATAAATCAATGAATCGCTGTCGACACTG 430

Qy 491 ACGCTTTTAAAGTGACGGTAAATGCGCGAGAAATTTGAATTAACAAAAACAGAGTATGATT 550
Db 431 CCGCCACAAAAGTGTGTTTGTAGAAAATGAAGTTATTTTACAAGCAAGAATACAAAT 490

Qy 551 TACTATATCTTCTAGCTGAAAAATAAAACCATGTTATGCAACGGGAACAAATTTTAAATC 610
Db 491 TGCTGGTTTCTTGTATGGAACATCCTAACCGTGTGTGGAATAAAGAGAAGAACTTTTGAAA 550

Qy 611 ATGTATGGGGTTATAAT 627
Db 551 GTGTCTGGGGTTTGAT 567
```

Search completed: October 5, 2004, 09:25:31
Job time : 173.368 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 5, 2004, 02:32:01 ; Search time 6114.39 Seconds
(without alignments)
10749.507 Million cell updates/sec

Title: US-09-006-627-1
Perfect score: 2201
Sequence: 1 TAATTTAAAGCAACTATT.....AAAACAAATAACAGTCGTT 2201

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues
Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST.*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estm:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_est3:*
12: gb_est4:*
13: gb_est5:*
14: gb_est6:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	111	5.0	1200	13 BX437758	BX437758 BX437758
2	107	4.9	1626	14 CF238805	CF238805 AGENCOURT
c 3	93.2	4.2	1392	29 CG757503	CG757503 P052-4-CO
4	91.6	4.2	1201	9 AL565455	AL565455 AL565455

c 5	91	4.1	1252	29	CG755566	CG755566 P051-1-F0
c 6	91	4.1	1348	29	CG749499	CG749499 P043-4-A0
c 7	89.4	4.1	1200	13	BX437739	BX437739 BX437739
c 8	87.4	4.0	994	13	BX414650	BX414650 BX414650
c 9	86.2	3.9	1074	28	BZ569636	BZ569636 SP_BA009
c 10	86.2	3.9	1566	29	CG757757	CG757757 P053-1-D0
c 11	86	3.9	1135	29	CNS033CQ	AL226115 Tetradon
c 12	84.8	3.9	1491	29	CG753221	CG753221 P048-2-A0
c 13	84.6	3.8	1811	29	CG753732	CG753732 P048-4-G0
c 14	84	3.8	1045	13	BX456814	BX456814 BX456814
c 15	83.8	3.8	1376	29	CG747831	CG747831 P041-3-B0
c 16	83.6	3.8	853	29	CG767791	CG767791 TC848.2C
c 17	83.6	3.8	1101	29	CNS0039G	AL063921 Drosophill
c 18	83.4	3.8	1313	29	CG751144	CG751144 P045-3-E0
c 19	83.4	3.8	1454	29	CG747614	CG747614 P041-2-A0
c 20	82.6	3.8	1373	29	CG750869	CG750869 P045-2-E1
c 21	82.4	3.7	1896	29	CG753083	CG753083 P048-1-C0
c 22	82.2	3.7	886	29	CNS01JEQ	AL146947 Anopheles
c 23	82.2	3.7	1101	29	CNS00EVL	AL069706 Drosophill
c 24	81.8	3.7	1353	29	CG744812	CG744812 P037-3-B0
c 25	81.4	3.7	1210	29	CG749728	CG749728 P044-1-C0
c 26	81	3.7	1174	29	CG746305	CG746305 P039-3-A0
c 27	80.8	3.7	1324	29	CG746828	CG746828 P040-1-G1
c 28	80.6	3.7	1361	29	CG744327	CG744327 P036-4-E0
c 29	80.4	3.7	1531	29	CG748014	CG748014 P041-4-B0
c 30	79	3.6	1201	13	BX439779	BX439779 BX439779
c 31	78.8	3.6	1201	13	BX424465	BX424465 BX424465
c 32	78.8	3.6	1269	29	CG757211	CG757211 P052-2-G0
c 33	78.8	3.6	1298	29	CG756607	CG756607 P051-4-C0
c 34	78.2	3.6	1364	29	CG757986	CG757986 P053-2-D0
c 35	78	3.5	949	14	CA465606	CA465606 AGENCOURT
c 36	78	3.5	1362	29	CG757918	CG757918 P053-2-A0
c 37	77.8	3.5	1200	13	BX415878	BX415878 BX415878
c 38	77.8	3.5	1221	28	CC301561	CC301561 CH261-13K
c 39	77.8	3.5	1433	29	CG745119	CG745119 P037-4-G0
c 40	77.4	3.5	1305	29	CG744200	CG744200 P036-3-H0
c 41	77.2	3.5	974	29	CNS00ITT	AL075432 Drosophill
c 42	76.8	3.5	1296	29	CG744840	CG744840 P037-3-C0
c 43	76.8	3.5	1331	29	CG749578	CG749578 P043-4-D1
c 44	76.6	3.5	1249	28	BZ566650	BZ566650 PAC82-164
c 45	76.4	3.5	1528	29	CG753854	CG753854 P049-1-D0

ALIGNMENTS

RESULT 1
BX437758
LOCUS
DEFINITION BX437758 Homo sapiens THYMUS Homo sapiens cDNA clone CS0CAP008VB01
5-PRIME, mRNA sequence.
ACCESSION BX437758
KEYWORDS BX437758.1 GI:30773605
SOURCE EST.
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1200)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Paradise Avenue Genoscope sequence ID : CS0CAP008CA01QPI.
Location/Qualifiers
1. .1200
/organism="Homo sapiens"

FEATURES
source

[illegible]

ACCESSION CG755566
VERSION CG755566.1 GI:37982189
KEYWORDS GSS.
SOURCE Pristionchus pacificus
ORGANISM Pristionchus pacificus
Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
Neodiplogasteridae; Pristionchus.
REFERENCE 1 (bases 1 to 1252)
AUTHORS Srinivasan,J., Sinz,W., Jesse,T., Wiggers-Perebolte,L., Jansen,K.,
Buntjer,J., van der Meulen,M. and Sommer,R.J.
TITLE An integrated physical and genetic map of the nematode Pristionchus
pacificus
JOURNAL Mol. Genet. Genomics 269 (5), 715-722 (2003)
MEDLINE 22835951
PUBMED 12884007
COMMENT Contact: Sommer RJ
Evolutionary Biology
Max-Planck-Institute for Developmental Biology
Spemannstr. 37-39, Tuebingen D-72076, Germany
Tel: 00497071601371
Fax: 00497071601498
Email: ralf.sommer@tuebingen.mpg.de
Class: BAC ends.
Location/Qualifiers
1..1252
/organism="Pristionchus pacificus"
/mol_type="genomic DNA"
/strain="California"
/db_xref="taxon:54126"
/clone_lib="Ppa EcoRI BAC Library"
/note="The library was generated by a partial digest of
the genomic DNA with EcoRI and cloning into the BAC
vector."

ORIGIN
Query Match 4.1%; Score 91; DB 29; Length 1252;
Best Local Similarity 51.1%; Pred. No. 8e-05;
Matches 214; Conservative 0; Mismatches 205; Indels 0; Gaps 0;
QY 74 TATGAGGTGTACAAATGACGCAATTTTAAATAGTAGAGATCAACAAACCTTAGCAAGAT 133
DB 462 TATGCAATATATTTTCATGAAGATTTTACTTATTGAAGATTAATCAAGGCCAGGAGT 403
QY 134 TTCTTGAATTTGAATCTACACATGACAAATTTACAATTTGACACAGATGATGACAAAG 193
DB 402 GGGTAACGACGGGGTTCCGAAGCGGGTTATGTCATTGATCGCGTTCTGATGSCAGAG 343
QY 194 ACGGTTTATGATAAGCGGTTAGCCATTTACTATGATTTTATCATATATGATTTAATGTTGC 253
DB 342 ATGGGCTTTATCTTCGCTGGAAGGATGATTTATGCAATGATTCATCTCGGATATTATGCTTC 283
QY 254 CGTCAATTAATGGCTTTAGAAATTTCTCGCAAAATTTAGACAAACCAATCTACACCTATCA 313
DB 282 CGGTTATGATGGCTGGCGAGATCTTACAAACGTTAAGAACACCAAGCAACCCCTGTTA 223
QY 314 TTATTAATTAACAGCAAAAGTGTATGATGACAAAGTTGCTGGGCTTGATACGGTGCAG 373
DB 222 TTGCTCTTACTCAAGGGATTTCTGCGATGACAGAGTCAGAGGCTGGACAGTGGAGCAA 163
QY 374 ACGATTATATAGTTAAGCGGTTTGATATTTAGAAAGCTTTTAGCAAGAATTCGTGCAATTT 433
DB 162 ATGATTATCTGTGTAAGACCTTTTTCATTTTCTGAGTTGCTGCAAGGGTTCGGGCACAAAT 103
QY 434 TACGTCGTGACGACCAAAAGGATATTATCGATGTCACAGGTTATCAATTTGATAGAAC 492
DB 102 TAAGCAACATCACGCTTTGAATTCGAGCTCGGTACCCGGGATCCTCTAGAGTGCACC 44

RESULT 6
CG749499/c
LOCUS P043-4-A06.za Ppa EcoRI BAC Library Pristionchus pacificus genomic,
DEFINITION genomic survey sequence.
1348 bp DNA linear GSS 24-OCT-2003

ACCESSION CG749499
VERSION CG749499.1 GI:37970425
KEYWORDS GSS.
SOURCE Pristionchus pacificus
ORGANISM Pristionchus pacificus
Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
Neodiplogasteridae; Pristionchus.
REFERENCE 1 (bases 1 to 1348)
AUTHORS Srinivasan,J., Sinz,W., Jesse,T., Wiggers-Perebolte,L., Jansen,K.,
Buntjer,J., van der Meulen,M. and Sommer,R.J.
TITLE An integrated physical and genetic map of the nematode Pristionchus
pacificus
JOURNAL Mol. Genet. Genomics 269 (5), 715-722 (2003)
MEDLINE 22835951
PUBMED 12884007
COMMENT Contact: Sommer RJ
Evolutionary Biology
Max-Planck-Institute for Developmental Biology
Spemannstr. 37-39, Tuebingen D-72076, Germany
Tel: 00497071601371
Fax: 00497071601498
Email: ralf.sommer@tuebingen.mpg.de
Class: BAC ends.
Location/Qualifiers
1..1348
/organism="Pristionchus pacificus"
/mol_type="genomic DNA"
/strain="California"
/db_xref="taxon:54126"
/clone_lib="Ppa EcoRI BAC Library"
/note="The library was generated by a partial digest of
the genomic DNA with EcoRI and cloning into the BAC
vector."

ORIGIN
Query Match 4.1%; Score 91; DB 29; Length 1348;
Best Local Similarity 43.7%; Pred. No. 7.7e-05;
Matches 487; Conservative 0; Mismatches 621; Indels 7; Gaps 2;
QY 1080 ATAAAAAACCGCTATAAGCGCATTCGAATTTTAAATTTAAAGAACCAATTTACACGGAA 1139
DB 1335 AAATATATAAATTTATATAAATTTAAATTTAAATATAAATAAATAAATAAATAAATAT 1276
QY 1140 GATTTCAAAGGGTATAGCTTGTAAATTCATTCACCTAGAAAAATTTATGATAACATCGTAAAA 1199
DB 1275 TATATATAAATAATAATAAATTTATATTTATTTATAAATAAATAAATAAATAAATAA 1216
QY 1200 TCATTGTATATCATTGGCTGGCATTTGGAGTGATTTGCCAACAAATTTAACTGCCACAATC 1259
DB 1215 AAATTTATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 1156
QY 1260 AGTTATGTATTTTCAACACAAATTTACTAAACCGCTTGTTCAGTTTATCAAAATAAATGATT 1319
DB 1155 AATATTTTAAATAAATTTTATATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 1096
QY 1320 GAGATTGACGAGATGGTTTTTCAAAATAAATTTGCAATTTAAATACAAATTTATCAAGAAATA 1379
DB 1095 AAATATATAAATAAATTTATATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 1036
QY 1380 GATATTTAGCAAAATAGCTTTTAAATGAGATGATGAGCCAAATTTGAAGAATTCATTTAATCAA 1439
DB 1035 TATAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 976
QY 1440 CAAGACAAATTTGT-TGAGATGCGTCACATGATTTACGACACCATTTACAAATTTATTTCA 1498
DB 975 AAAAAATTTTAAATTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 916
QY 1499 AGGTCATTTAAATTTGATTTTCAGCGATGGGAAAAAAGACCCAGCAGTATTAGAGAATC 1558
DB 915 AATAAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 856
QY 1559 GTTAAATTTTCTATTGAGAAATGAATCGTATCATATAAATTTAGTCGAGAAATTTACTTGA 1618

```

Db 855 AATAAAAAATAAAAAATAAAAAATTTAAATAAAAATATAATATAAATAAATAA 796
Qy 1619 ATTGACTAAAGGAGATGTABATGACATTTCTTCTGAAGCGCAGACCGTGCATATTATGA 1678
Db 795 AAAAAAATATAAATAAATAAATAAATAATATTATAAATAATAAATAAATAAATAAATAA 736
Qy 1679 TGAATATCGCTCGGGAATACACTCATTTAAACAAATTTGCATCCTGATTAATCAATTTGATAC 1738
Db 735 TTAATAAANAATAAATAAATAAATAAATAATATAATATAAATAAATAAATAAATAAATAA 676
Qy 1739 GGATCTGCATCTAAATAATCTAGAAATTAATAATGAACCTCATCAATTCGACCAATATT 1798
Db 675 ATATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 622
Qy 1799 TTTAATCTTTATTCATATGCAATCAATATGATGTGAAGATAAGAAATTAAGATTAA 1858
Db 621 AATATTTATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 562
Qy 1859 GACAAGGTTAAAAATTAAGCAAAATAAATTAAGAAATTAACAGATCATCGAATTTGATATCC 1918
Db 561 AAAAAATNTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 502
Qy 1919 AGAGGAAGATCAAGATTTTCATTTTGTATCGCTTTTATCGAGTGATAAATCTCGTCAAG 1978
Db 501 AAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 442
Qy 1979 AAGTCAAGCGGTAATGACTCGGATTTATCTATTGCTCAAAATATCAATTTAAACGG 2038
Db 441 TAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 382
Qy 2039 AGGATCGATTAAAAATTAAGTGAAATTAATAAGGAACACGTTTAAAAATCATATTTTA 2098
Db 381 TTTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 322
Qy 2099 ATCATGTCTGAGCGTCAATCAAAAGTCATAGATCAATTTTTTAAAGTACATTTAGCTGT 2158
Db 321 ATTATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 262
Qy 2159 GACTAATGTATAAGCAACTATATAAACAATAA 2193
Db 261 AAAAAATTTATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 227

```

```

RESULT 7
BX437739/c
LOCUS
DEFINITION BX437739 Homo sapiens THYMUS Homo sapiens cDNA clone CSOCAP008YG24
3-PRIME, mRNA sequence.
BX437739
ACCESSION BX437739.1 GI:31018315
VERSION
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1200)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 534.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CSOCAP008BD12NP1&cluster=534.r. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CSOCAP008BD12NP1.
Location/Qualifiers
1. 1200
FEATURES
source

```

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSOCAP008YG24"
/tissue_type="THYMUS"
/clone_lib="Homo sapiens THYMUS"
/notes="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."

```

ORIGIN

```

Query Match 4.1%; Score 89.4; DB 13; Length 1200;
Best Local Similarity 35.4%; Pred. No. 0.0004;
Matches 306; Conservative 120; Mismatches 434; Indels 4; Gaps 2;

Qy 1083 AAAAAACGCTATAAAGGCATTTGAATATTTAAATTTATTAAGAACCAATTAACAAGCGCAAGAT 1142
Db 1173 ATHHMMTHMTHTHWHHTTHMMWMTTITTTMMWMTTMMWMTTITTTTTHMMWMTHT 1114
Qy 1143 TTCAAAGGGTATAGCTTGTAAATTCATCTACAGAAAAATATGATACATCGTGAATAATCA 1202
Db 1113 TWTMTTMTTMTTCWTITTTTMTTHTTMMHTTMTTITTTMMWMTTMMWMTTMMWMTTMMWMTT 1054
Qy 1203 TTGTATATCATTCGCGCTGGCATTTGGAGTGATTTGCAACAATTTAACTGCCACATCAGT 1262
Db 1053 TWTMTTMTTMTTMTTMMWMTTATTTTHTACMTMMWMTTMMWMTTMMWMTTMMWMTTMMWMT 994
Qy 1263 TATGATTTTTCACACAAAA--TTACTAAACCGCTTGTCTAGTTTATCAAAATAAATGATTCG 1320
Db 993 THTTTTTTTTHAWAMAMTTTTTTHMTMAHMTMTTAAATTTTATANAATTTTTTAT 934
Qy 1321 AGATTCGACGAGATGGTTTTCAAAAATAAATTCGAATTAATACAAATATATGAAGAAATAG 1380
Db 933 AATTWATWATATATATATAATAATATATCAATTTTAAATTTTATATATTTTTTTT 874
Qy 1381 ATAAATTTAGCAATACGTTTAAATGAGATGATGAGCCAAATTTGAAGATCATTTTAAATCAAC 1440
Db 873 ATTTTATWTTTHTTATTTTATWTTTAAATTAATTTTATWTTTATWTTTATWTTTATWTTTAT 814
Qy 1441 AAAGAC--AATTTGTTGAAGATCGCTACATGAATATACGAACACCATTTACAAATTTATTC 1498
Db 813 TAATTTTATTTATATAAAMWMTTWTTTTAAATATATATATATAAATAATATAATAA 754
Qy 1499 AGTCAATTTAAATTTGATTCAGCGATGGGAAAAAAGAACCCAGCAGTATTAGAAGATC 1558
Db 753 AMTTHAAATTTATTTTATTTTCAAWAAAAACCAATAAATAAATAAATAAATAAATAAATAA 694
Qy 1559 GTTAAATATTTCTATTGAAGAAATGAATCGTATCATATAAATTTAGTCGAAGAAATTTACTTGA 1618
Db 693 AYAAAAATTTATTTTWWAAAYATWTTTATTTATTAATAAATAAATAAATAAATAAATAAATAA 634
Qy 1619 ATTGACTAAAGGAGATGTAAATGACATTTCTTCTGAAGCGCAGACCGTGCATATTATGA 1678
Db 633 WAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 574
Qy 1679 TGAATTCGCTCGCAATACACTCATTTAAACAAATTTGCATCTCGTATTCATCAATTTGATAC 1738
Db 573 HATAATATATAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 514
Qy 1739 GGATCTGCATCTAAAAATCTAGAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTT 1798
Db 513 AAAAAAATAAATCAATTAATTTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 454
Qy 1799 TTTAATCTTTTATGATAATGCAATCAAAATATGATGTGAAGAAATTAAGAAATTTAAAGTTAA 1858
Db 453 ATAATAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 394
Qy 1859 GACAAGGTTAAAAATAAGCAAAAAATAATTTGAATTTACAGATCATCGAATTTGGTATTC 1918
Db 393 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 334

```


[illegible]

RESULT 10	CG757757	1566 bp	DNA	linear	GSS 24-OCT-2003
LOCUS	P053-1-D07-2c Ppa EcoRI	BAC Library	Pristionchus pacificus	genomic,	
DEFINITION	genomic survey sequence.				
ACCESSION	CG757757				
VERSION	CG757757.1	GI:37986636			
KEYWORDS	GSS.				
SOURCE	Pristionchus pacificus				
ORGANISM	Pristionchus pacificus				
REFERENCE	Eukaryota: Metazoa; Nematoda; Chromadorea; Diplogasterida;				
AUTHORS	Srinivasan, J., Sinz, W., Jesse, T., Wiggers-Perebolte, L., Jansen, K., Buntjer, J., van der Meulen, M. and Sommer, R. J.				
TITLE	An integrated physical and genetic map of the nematode Pristionchus pacificus				
JOURNAL	Mol. Genet. Genomics 269 (5), 715-722 (2003)				
MEDLINE	22835951				
PUBMED	12884007				
COMMENT	Contact: Sommer RJ Evolutionary Biology Max-Planck-Institute for Developmental Biology Spemannstr. 37-39, Tuebingen D-72076, Germany Tel: 00497071601371 Fax: 00497071601498 Email: ralf.sommer@tuebingen.mpg.de Class: BAC ends.				
FEATURES	Location/Qualifiers				
source	1. .1566 /organism="Pristionchus pacificus" /mol_type="genomic DNA" /strain="California" /db_xref="taxon:54126" /clone_lib="Ppa EcoRI BAC Library" /note="The library was generated by a partial digest of the genomic DNA with EcoRI and cloning into the BAC vector."				
ORIGIN					
	Query Match 3.98; Score 86.2; DB 29; Length 1566; Best Local Similarity 36.44; Pred. No. 0.00039; Matches 445; Conservative 0; Mismatches 778; Indels 0; Gaps 0;				
Qy	973 AAGAGATAATTATTATGATGAGCATATATAAATTTATTGTGAGCATCGAATGATAACA	1032			
Db	1512 AA	1453			
Qy	1033 CAGTGAGAGTTGAACCCAGGTTATGAACACCGGTTATTTTGTGACCGGTTATTAACAGGCT	1092			
Db	1452 ANANANNN	1393			
Qy	1093 ATAAGGCATTGAATATTATTTATTTAAAGAAACCAATTACACGCAAGATTTCAAGGGT	1152			
Db	1392 AA	1333			
Qy	1153 ATAGCTTGTTAAATTCATTCTACTAGAAATTTATGATACATCGTAAATCATTTGTATATCA	1212			
Db	1332 AA	1273			
Qy	1213 TTGCGCTGGCATTGTTGGAGTGATTGCAACAAATATACTATTAACATCGCACAAATCAGTTATGTTT	1272			
Db	1272 AA	1213			
Qy	1273 CAACACAAATTAATAACCGCTTGTCAGTTTATCAATAATAATGATTGAGATTGACGAG	1332			
Db	1212 AA	1153			
Qy	1333 ATGCTTTTCAAAATTAATTCGAATTAATACAAATATTGAAGAAATAGATAATTTAGCAA	1392			
Db	1152 ANNNAN	1093			
Qy	1393 ATAGCTTTAATGAGATGATGAGGCAAAATTTGAAGAATCATTTTAATCAACAAGACAATTTG	1452			

```
Db 1092 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1033
Qy 1453 TTGAAGATGCGTCACATGAATTTACGAACACCATTTACAAATTTTCAAGTCATTTAAATTT 1512
Db 1032 ANAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 973
Qy 1513 TGATTCAGGATGGGGAAGAAAGACCCAGCAGTATTAGAAGATCGTTAAATATTCTTA 1572
Db 972 NAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 913
Qy 1573 TTGAAGATGATCGTATCAATAAATTTAGTCGAGAATTTACTTGAATGCTAAAGAG 1632
Db 912 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 853
Qy 1633 ATGTAATGACATTTCTTCTGAAGCGCAGACCGTCATATTAAATGATGAATTCGCTCGC 1692
Db 852 ANNNAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 793
Qy 1693 GAATACACTATTAACAAATTTGCATCTGATTTCAATTTGATACGATCTGACATCTA 1752
Db 792 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 733
Qy 1753 AAAATCTAGAATTAATAGTAACCTCATCAATTCGAACAATTTATTTAACTTTATTG 1812
Db 732 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 673
Qy 1813 ATAATGCAATCAATATGATGTGAAGATAAGAAATTAAGTTAAGCAAGTTAAAAA 1872
Db 672 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 613
Qy 1873 ATAAGCAAAAAATAATGAAATTTACAGATCATGGAATTTGGTATTCAGAGGAAGATCAAG 1932
Db 612 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 553
Qy 1933 ATTTCAATTTTGATCGCTTTTATCGAGTGGATAAATCTCGTTCAAGAAGTCAGCGGTA 1992
Db 552 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 493
Qy 1993 ATGGACTCGATTTATCTTGTCTCAAAATATCAATTTAAACGGAGATCGATTAAAA 2052
Db 492 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 433
Qy 2053 TTAAGATGAAATTAATAAGGAACAACGCTTTTAAATCATATTTTAAATCATGCTCTGAGAC 2112
Db 432 NAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 373
Qy 2113 GTCAATCAAGTCATAGGATCAATTTTAAAGTACACATGCTGCTGACTAATGATTAAG 2172
Db 372 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 313
Qy 2173 AACCACTATAAACAATAAACA 2195
Db 312 AAAAAAAAAAAAAAAAAAAAAAAAAA 290
```

```
RESULT 11
CNS033GQ 1135 bp DNA linear GSS 01-SEP-2000
LOCUS Tetraodon nigroviridis genome survey sequence PUC-ori end of clone
DEFINITION 208P24 of library G from Tetraodon nigroviridis, genomic survey
sequence.
ACCESSION AL226115
VERSION AL226115.1 GI:7885026
KEYWORDS GSS; genome survey sequence.
SOURCE Tetraodon nigroviridis
ORGANISM Tetraodon nigroviridis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodontidae; Tetraodon.
REFERENCE 1
AUTHORS Roest Crolius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,
```

```
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
Roest Crolius,H., Jaillon,O., Dasilva,C., Ozouf-Costaz,C.,
Fizames,C., Fischer,C., Bouneau,L., Billault,A., Quetier,F.,
Saurin,W., Bernot,A. and Weissenbach,J.
Estimate of human gene number provided by genome-wide analysis
using Tetraodon nigroviridis DNA sequence
Nat. Genet. 25 (2), 235-238 (2000)
20296633
10835645
2
Roest Crolius,H., Jaillon,O., Dasilva,C., Ozouf-Costaz,C.,
Fizames,C., Fischer,C., Bouneau,L., Billault,A., Quetier,F.,
Saurin,W., Bernot,A. and Weissenbach,J.
Characterization and repeat analysis of the compact genome of the
freshwater pufferfish Tetraodon nigroviridis
Genome Res. 10 (7), 939-949 (2000)
20359837
10899143
3 (bases 1 to 1135)
Genoscope.
Direct Submission
Submitted (12-APR-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 Evry cedex - FRANCE (E-mail : segref@genoscope.cns.fr)
- Web : www.genoscope.cns.fr)
This sequence is a single read and was generated as part of a large
scale clone-end sequencing project of the Tetraodon nigroviridis
genome. For more information, please take a look at
http://www.genoscope.cns.fr/Tetraodon.
FEATURES
Location/Qualifiers
1..1135
/organism="Tetraodon nigroviridis"
/mol_type="genomic DNA"
/db_xref="taxon:99883"
/clone="208P24"
/clone_lib="G"
/note="Genoscope sequence ID : COAG208DH12Sp1-end :
PUC-ori"
```

```
ORIGIN
Query Match 3.9%; Score 86; DB 29; Length 1135;
Best Local Similarity 42.8%; Pred. No. 0.00049;
Matches 394; Conservative 14; Mismatches 512; Indels 1; Gaps 1;
Qy 973 AAGAGATAATTTATTTATGATGAGCATATAATAAATTTATTTGAGACATCGAATGATAACA 1032
Db 215 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 274
Qy 1033 CAGTGAGAGTTGACCCAGGTTATGAACACCGTTATTTTGACCGGCTAATAAATAACGCT 1092
Db 275 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 334
Qy 1093 ATAAAGGCATTTGAATATTATTATTAAGAAACCAATTTACACGACAGATTTCAAAAGGT 1152
Db 335 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 394
Qy 1153 ATAGCTTTGTTAATTCATTCATAGAAAATTTATGATAACATCGTAAATCATGTTATATCA 1212
Db 395 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 454
Qy 1213 TTGCGCTGGCATTTGGAGTGATTCGCAACATTTATACTGCCACATCGTTATGTTATTTT 1272
Db 455 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 514
Qy 1273 CAACACAAATTTACTTAAACCGCTGTCAGTTTATCAAAATTAATGATGAGATTCGACGAG 1332
Db 515 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 574
Qy 1333 ATGCTTTTCAAAATAAATTTGCAATTTAAATCAAAATTTATGAAGAATAGATTAATTTAGCAA 1392
Db 575 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 634
Qy 1393 ATAGCTTTAATGAGATGATGCGCAATTTGAAGATCATTTTAATCAACAAGACATTTG 1452
Db 635 AAAAAATWAAAAAATAAATAAATTTTAAAAAATAAATTAATTTTAAAAAATAAATWATWTA 694
```



```

BX456814/c      BX456814      1045 bp      mRNA      linear      EST 22-MAY-2003
LOCUS           BX456814      Homo sapiens THYMUS Homo sapiens cDNA clone CS0CAP004YO19
DEFINITION      3-PRIME, mRNA sequence.
ACCESSION       BX456814
VERSION         BX456814.1 GI:31034806
KEYWORDS        EST.
SOURCE          Homo sapiens (human)
ORGANISM        Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE       1 (bases 1 to 1045)
AUTHORS        Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE          Full-length cDNA libraries and normalization
JOURNAL         Unpublished (2001)
COMMENT        Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 86.f. For more
information about this cluster, see http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0CAP004AH10NP1&cluster=86.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paradise Avenue Genoscope sequence ID : CS0CAP004AH10NP1.
FEATURES
    source
    Location/Qualifiers
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /clone="CS0CAP004YO19"
        /tissue_type="THYMUS"
        /clone_lib="Homo sapiens THYMUS"
        /note="vector: pCMVSPORT 6; 1st strand cDNA was primed
        with a NotI-oligo(dT) primer. Five prime end enriched,
        double-strand cDNA was digested with Not I and cloned into
        the Not I and EcoRV sites of the pCMVSPORT 6 vector.
        Library was not normalized."
ORIGIN
Query Match          3.8%; Score 84; DB 13; Length 1045;
Best Local Similarity 45.2%; Pred. No. 0.001;
Matches 299; Conservative 31; Mismatches 318; Indels 13; Gaps 3;
Qy 1238 AACATTATACGCCACATCAGTTATGTATTTTCAACACAAATTTACTAACCGCTTGT 1297
Db 1000 AWAATTAWAATWAAAAAATAATTAATWAATTTTATAATATTTTAAATTAATTAATTT 941
Qy 1298 CAGTTTATCAATTAATGATTCGATTCGACGAGATGGTTTTCAAAATTAATGCAATT 1357
Db 940 TTAATAATTAATTTWATTTATTTAAATTAATAAATAAATTTTMAWATTTTAAATTTT 881
Qy 1358 AAATACAAATATGAAGAAATAGATAATTTAGCAATACGTTTAAATGAGATGAT-GAGCC 1416
Db 880 TTAAWAAWAAATATAATTAATTTTATAATTAATTAATTTTAAATTAATTAATTAAT 821
Qy 1417 AAATGAAGATCATTTTAAATCAACAAACAAATTTGTGAAGATGCGTCATCAATTTAC 1476
Db 820 AAATTTTAAAAAATAATTAATAAATAATTAATTAATTAATTAATTAATTAATTAAT 761
Qy 1477 GAACACCATTAACAATTTTACAGTTCATTTAAATTTGATTCAGCGATGGGAAAAAAG 1536
Db 760 AATAATAAATAAATAATTAATTAATAAATAAATAATTAATTAATTAATTAATTAAT 701
Qy 1537 ACCCAGCAGTATTAGAAGAAATCGTTAAATATTTCTATTGAAGAAATGAATCGTATCAAA 1596
Db 700 TAAAAATTTTAAAAATTAATTAATTTTAAATTTTAAATTTTAAATTAATTAATTAAT 641
Qy 1597 AATTAGTCGAAGAAATTTACTTGAATT--GACTAAAGGAGATGTAAATGACATTTCTCTGA 1654
Db 640 AATAAAAATAAAATATAATTAATAAATAAATAATTAATTTTAAATTAATTAATTTTAA 581

```

```

Qy 1655 AGCGCAGACGCGTCATATTAAATGATGAAATTCGCTCGCAATACACTCATTTAAACAATT 1714
Db 580 AATATATTATTTTATTATTAATAATTTATTTT-----TATATTAATAATAAAATTT 531
Qy 1715 GCATCCTGATTATCAATTTTGATACGGATCTGACATCTAAAAATCTAGAAATGAA 1774
Db 530 TTTTAAATTAATATTTTWTAAATTAATATTTTATATAAATAATTTTAAATAATATATTT 471
Qy 1775 ACCTCATCAATTCGAACAATTTATTTTAATCTTTTATTTGATATGCAATCAATATATGAT 1834
Db 470 AATTTATTATTAATAAAAAATTTTTTTAAATTAATTTTAAATTTTAAATTTTAAATTTAT 411
Qy 1835 GAAGATTAAGAAATTAAGTTAAGCAAGGTTTAAAAAATAAGCAAAAAATAATTCGAAT 1894
Db 410 AAAYAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATA 351
Qy 1895 T 1895
Db 350 Y 350
RESULT 15
CG747831/c
LOCUS           CG747831
DEFINITION      1376 bp      DNA      linear      GSS 24-OCT-2003
ACCESSION       CG747831
VERSION         CG747831.1 GI:37968757
KEYWORDS        genomic survey sequence.
SOURCE          Pristionchus pacificus
ORGANISM        Pristionchus pacificus
Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
Neodiplogasteridae; Pristionchus.
REFERENCE       1 (bases 1 to 1376)
AUTHORS        Srinivasan, J., Sinz, W., Jesse, T., Wiggers-Perebolte, L., Jansen, K.,
Buntjer, J., van der Meulen, M. and Sommer, R.J.
TITLE          An integrated physical and genetic map of the nematode Pristionchus
pacificus
JOURNAL         Mol. Genet. Genomics 269 (5), 715-722 (2003)
MEDLINE         22835951
PUBMED          12884007
COMMENT        Contact: Sommer RJ
Evolutionary Biology
Max-Planck-Institute for Developmental Biology
Spemannstr. 37-39, Tuebingen D-72076, Germany
Tel: 00497071601371
Fax: 00497071601498
Email: ralf.sommer@tuebingen.mpg.de
Class: BAC ends.
FEATURES
    source
    Location/Qualifiers
        1..1376
        /organism="Pristionchus pacificus"
        /mol_type="genomic DNA"
        /strain="California"
        /db_xref="taxon:54126"
        /clone_lib="ppa EcoRI BAC Library"
        /note="The library was generated by a partial digest of
        the genomic DNA with EcoRI and cloning into the BAC
        vector."
ORIGIN
Query Match          3.8%; Score 83.8; DB 29; Length 1376;
Best Local Similarity 39.2%; Pred. No. 0.00097;
Matches 361; Conservative 0; Mismatches 561; Indels 0; Gaps 0;
Qy 1274 AACACAAATTAATAACCGCTTGTTCAGTTTATCAATAAATAATGAGATTCGACGAGA 1333
Db 1327 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1268
Qy 1334 TGGTTTCAAAATAAATTCGAATTAATAACAATTAATTAAGAAATAGATAATTTAGCAA 1393
Db 1267 ANAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 1208

```

```
Qy 1394 TACGTTTAAATGAGATGATGAGCCAAATTTGAGAAATCATTTTAATCAACAAAGACAAATTTGT 1453
Db 1207 AAAAAAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANA 1148
Qy 1454 TGAAGATCGTCACATGATTAAGAACACCACTTACAAAATTTTCAAGGCTCATTTAAATTT 1513
Db 1147 AAAAAAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANA 1088
Qy 1514 GATTCAGCGATGGGGAAGAAAAAGCCAGCAGTATTAGAGAATCGTTTAAATATTTCAT 1573
Db 1087 NAAAAAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANA 1028
Qy 1574 TGAAGAAATGAATCGTATCATTAATTAATTAAGTCGAGAAATTAATTAAGTAAAGGAGA 1633
Db 1027 NAAAAAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANA 968
Qy 1634 TGTAATGACATTTCTTCTGAGCGCAGACCGTGCATATTAAATGATGAATTCGCTCGCG 1693
Db 967 ANAAAAAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANA 908
Qy 1694 AATACACTCATTAACAATTCGATCTGATTTATCAATTTTGATACGGATCTGACATCTAA 1753
Db 907 NAAAAAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANA 848
Qy 1754 AAATCTAGAAATTAATTAAGAACCTCATCAATTCGAACAATTTTAAATCTTTATTGA 1813
Db 847 AAAAAAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANA 788
Qy 1814 TAATGCAATCAATATGATGTGAGAAATTAAGAAATTAAGTTAAGACAAGGTTAAAAAA 1873
Db 787 AAAAAAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANA 728
Qy 1874 TAAGCAAAAAATTAATTAAGAAATTAAGATCATGGAATTCGATTCAGAGGAAGATCAAGA 1933
Db 727 AAAAAAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANA 668
Qy 1934 TTTCAATTTTGTGCTGCTTTTATCGAGTGGATAAATCTCGTTCAAGAACTCAAGCGGTAA 1993
Db 667 AAAAAAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANA 608
Qy 1994 TGGACTCGGATTTATCTATTGCTCAAAAAATCATTCATTAACGCGAGGATCGATTAATAT 2053
Db 607 AAAAAAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANA 548
Qy 2054 TAAAGTGAAATTAATAAGGAACAACGTTTAAATCATATTTTAAATCATGCTCGAGACG 2113
Db 547 AAAAAAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANA 488
Qy 2114 TCAATCAAGTCATAGGATCAATTTTAAAGTACACATTAAGTGTGACTAATGTATAGA 2173
Db 487 AAAAAAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANA 428
Qy 2174 ACNACTATAAACCAATTAACA 2195
Db 427 NAAAAAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANA 406
```

Search completed: October 5, 2004, 09:21:44
Job time : 6123.39 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 4, 2004, 18:03:18 ; Search time 16.0164 Seconds
(without alignments)
705.906 Million cell updates/sec

Title: US-09-006-627-4

Perfect score: 1122

Sequence: 1 MTQILIVEDQNLFARFLELE.....KPYDRDXMIETVRGVGYVIR 219

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pap:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pap:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pap:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pap:*
5: /cgn2_6/ptodata/2/iaa/PTUS_COMB.pap:*
6: /cgn2_6/ptodata/2/iaa/backfiles.pap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1122	100.0	219	3	US-08-771-110-2
2	1122	100.0	219	4	US-09-642-000-4
3	968	86.3	222	4	US-09-134-001C-3212
4	654	58.3	228	4	US-09-107-532A-6536
5	468.5	41.8	234	4	US-09-634-238-220
6	458	40.8	246	4	US-09-107-532A-4135
7	452	40.3	246	4	US-09-134-001C-3175
8	449.5	40.1	234	4	US-09-107-532A-6746
9	449	40.0	233	4	US-09-082-077-4
10	449	40.0	234	4	US-09-134-001C-3063
11	449	40.0	239	4	US-08-956-171E-5239
12	442	39.4	234	1	US-08-850-118-2
13	442	39.4	234	2	US-09-008-253-2
14	442	39.4	234	3	US-09-093-335-2
15	442	39.4	245	4	US-09-134-001C-3779
16	435.5	38.8	245	4	US-09-252-991A-31392
17	435	38.8	184	4	US-09-134-000C-3820
18	432.5	38.5	234	3	US-08-771-321-2
19	430.5	38.4	231	4	US-09-134-000C-5664
20	430.5	38.4	239	4	US-09-711-164-367
21	429	38.2	227	4	US-09-328-352-8175
22	425	37.9	231	2	US-08-286-819A-12
23	425	37.9	231	3	US-08-286-819A-43
24	425	37.9	231	3	US-08-980-357-12
25	425	37.9	231	3	US-08-980-357-43
26	425	37.9	2296	2	US-08-286-819A-27
27	425	37.9	2296	3	US-08-980-357-27

Sequence 6135, Ap
Sequence 26894, A
Sequence 4846, Ap
Sequence 10454, A
Sequence 6873, Ap
Sequence 5145, Ap
Sequence 24389, A
Sequence 25108, A
Sequence 222, Appl
Sequence 45, Appl
Sequence 20963, A
Sequence 2, Appl
Sequence 4690, Ap
Sequence 13339, A
Sequence 6769, Ap
Sequence 30493, A
Sequence 11254, A

28 424 37.8 230 4 US-09-543-681A-6135
29 421 37.5 267 4 US-09-252-991A-26894
30 419.5 37.4 241 4 US-09-107-532A-4846
31 419 37.3 264 4 US-09-489-039A-10454
32 418.5 37.3 236 4 US-09-107-532A-6873
33 416.5 37.1 231 4 US-09-134-000C-5145
34 413.5 36.9 229 4 US-09-252-991A-24389
35 413 36.8 244 4 US-09-252-991A-25108
36 410.5 36.6 269 4 US-09-634-238-222
37 406 36.2 229 2 US-08-286-819A-45
38 406 36.2 229 3 US-08-980-357-45
39 404 36.0 230 4 US-09-252-991A-20963
40 403.5 36.0 243 2 US-08-771-455-2
41 401.5 35.8 249 4 US-09-328-352-4690
42 399 35.6 230 4 US-09-489-039A-13339
43 397.5 35.4 227 4 US-09-107-532A-6769
44 394.5 35.2 237 4 US-09-252-991A-30493
45 394 35.1 236 4 US-09-489-039A-11254

ALIGNMENTS

RESULT 1
US-08-771-110-2
; Sequence 2, Application US/08771110
; Patent No. 6084086
; GENERAL INFORMATION:
; APPLICANT: Wallis, Nicola G.
; APPLICANT: Hodgson, John E.
; TITLE OF INVENTION: No. 6084086el Compounds
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-2799
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/771,110
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Gimmi, Edward R.
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: P31316
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 219 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-771-110-2

Query Match 100.0%; Score 1122; DB 3; Length 219;
Best Local Similarity 100.0%; Pred. No. 5.1e-111;
Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MTQILIVEDQNLFARFLELELTHENYNVDYDQDGLDKALSHYYDLIILDLMLPSTING 60
Db 1 MTQILIVEDQNLFARFLELELTHENYNVDYDQDGLDKALSHYYDLIILDLMLPSTING 60
Qy 61 LEICRKIRQQOSTPIIIITAKSDTYDKVAGLDYGADDDYIVKPFDELLARIRAILRRQP 120

Db 61 LEICRKIRQOOSTPIIIITAKSDTYDKVAGLDYGADDDYIVKPFDEIELLARAILRRQP 120
QY 121 QKDIIDVNGITIDKNAFKVTNGAEIELTKTEYDLYLLAENKQHVQREQLNHVWGYN 180
Db 121 QKDIIDVNGITIDKNAFKVTNGAEIELTKTEYDLYLLAENKQHVQREQLNHVWGYN 180
QY 181 SEVETNVVDVYIRYLNRKLKPYDRDKMIETVRGVGVIR 219
Db 181 SEVETNVVDVYIRYLNRKLKPYDRDKMIETVRGVGVIR 219

RESULT 2

US-09-642-000-4
; Sequence 4, Application US/09642000
; Patent No. 6548281
; GENERAL INFORMATION:
; APPLICANT: Wallis, Nicola G.
; Shilling, Lisa K.
; Mooney, Jeffrey L.
; Debouck, Christine
; Zhong, Yi Yi
; Jaworski, Deborah D.
; Wang, Min
; Throup, John P.
; TITLE OF INVENTION: Histidine Kinase
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dechert, Price & Rhoads
; STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103-2793
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/642,000
; FILING DATE: 18-Aug-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/006,627
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Falk, Stephen T
; REGISTRATION NUMBER: 36,795
; REFERENCE/DOCKET NUMBER: GM10127
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-994-2488
; TELEFAX: 215-994-2222
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 219 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-642-000-4

Query Match 100.0%; Score 1122; DB 4; Length 219;
Best Local Similarity 100.0%; Pred. No. 5.1e-111;
Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MTQILIVDEQNLALELELTHENVNVDTEYDQGLDKALSHYVDLIILDLMLPSING 60
Db 1 MTQILIVDEQNLALELTHENVNVDTEYDQGLDKALSHYVDLIILDLMLPSING 60
QY 61 LEICRKIRQOOSTPIIIITAKSDTYDKVAGLDYGADDDYIVKPFDEIELLARAILRRQP 120

Db 61 LEICRKIRQOOSTPIIIITAKSDTYDKVAGLDYGADDDYIVKPFDEIELLARAILRRQP 120
QY 121 QKDIIDVNGITIDKNAFKVTNGAEIELTKTEYDLYLLAENKQHVQREQLNHVWGYN 180
Db 121 QKDIIDVNGITIDKNAFKVTNGAEIELTKTEYDLYLLAENKQHVQREQLNHVWGYN 180
QY 181 SEVETNVVDVYIRYLNRKLKPYDRDKMIETVRGVGVIR 219
Db 181 SEVETNVVDVYIRYLNRKLKPYDRDKMIETVRGVGVIR 219

RESULT 3

US-09-134-001C-3212
; Sequence 3212, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 3212
; LENGTH: 222
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3212

Query Match 86.3%; Score 968; DB 4; Length 222;
Best Local Similarity 84.5%; Pred. No. 1.1e-94;
Matches 185; Conservative 21; Mismatches 13; Indels 0; Gaps 0;

QY 1 MTQILIVDEQNLALELELTHENVNVDTEYDQGLDKALSHYVDLIILDLMLPSING 60
Db 4 MTNLIIVDEQNLALELTHENVNVDTEYDQGLDKALSKPYDLYLDMLPNG 63
QY 61 LEICRKIRQOOSTPIIIITAKSDTYDKVAGLDYGADDDYIVKPFDEIELLARAILRRQP 120
Db 64 LEICRKIRQOOSTPIIIITAKSDTYDKVAGLDYGADDDYIVKPFDEIELLARAILRRQP 123
QY 121 QKDIIDVNGITIDKNAFKVTNGAEIELTKTEYDLYLLAENKQHVQREQLNHVWGYN 180
Db 124 DKDVLIDGIIIDKDAFKVTNGHQLTKTEYDLYLLAENRNHVQREQLDHWGYN 183
QY 181 SEVETNVVDVYIRYLNRKLKPYDRDKMIETVRGVGVIR 219
Db 184 SEVETNVVDVYIRYLNRKLKPYDRDKMIETVRGVGVIR 222

RESULT 4

US-09-107-532A-6536
; Sequence 6536, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD/ROM ISO9660
; COMPUTER: PC

```
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 6536:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 228 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...228
; SEQUENCE DESCRIPTION: SEQ ID NO: 6536:
US-09-107-532A-6536

Query Match      58.3%; Score 654; DB 4; Length 228;
Best Local Similarity 56.4%; Pred. No. 2.2e-61;
Matches 133; Conservative 35; Mismatches 42; Indels 26; Gaps 4;

QY 1 MTQILVDEONLARFLELELTHENVVDTEYDQDGLDKALSHYDYLIIIDMLPSNG 60
DB 1 MSNLIIDEKNLARFVLELKHGEGYAEVHYNGRTGLDAAALNNEWDAILLDLMLPELNG 60

QY 61 LEICRKIRQOOSTPIIIITAKSDTYKVAGLDYGADDYIVKPFDEELLARIRAILRQP 120
DB 61 LEVCRRQVKNPIIMMTARSDVIRVSGLOHGADDYIVKPFDEELLARIRAILR-- 118

QY 121 QKDIIVNGGITDKNAFKVTV-----NGAE-IELTKTEYDLYLLAENK 163
DB 119 ----IDEG---DKNVAQTITRYDLTIKENRVVRGNEVIELTKREYELLTLMEV 171

QY 164 NHVMOREQILNHVWGYNSEVETNVVDVYIRYLNRKLPKPYDRDKMIETVRGVYIR 219
DB 172 NVVLARDVLLNKVMGYETEVETNVVDVYIRYLNRKIDVPGESYIQTIVRGTYVMR 227

RESULT 5
US-09-634-238-220
; Sequence 220, Application US/09634238
; Patent No. 6544772
; GENERAL INFORMATION:
; APPLICANT: Glenn, Matthew
; APPLICANT: Havukkala, Ilkka J.
; APPLICANT: Bloksberg, Leonard, N.
; APPLICANT: Lubbers, Mark W.
; APPLICANT: Dekker, James
; APPLICANT: Christensson, Anna C.
; APPLICANT: Holland, Ross
; APPLICANT: O'Toole, Paul W.
; APPLICANT: Reid, Julian R.
; APPLICANT: Coolbear, Timothy
; TITLE OF INVENTION: Polynucleotides, materials incorporating
; TITLE OF INVENTION: them and methods for using them.
; FILE REFERENCE: 11000.1043U1
; CURRENT APPLICATION NUMBER: US/09/634,238
; CURRENT FILING DATE: 2000-08-08

; NUMBER OF SEQ ID NOS: 422
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 220
; LENGTH: 234
; TYPE: PRT
; ORGANISM: Lactobacillus rhamnosus
US-09-634-238-220

Query Match      41.8%; Score 468.5; DB 4; Length 234;
Best Local Similarity 43.9%; Pred. No. 1.1e-41;
Matches 100; Conservative 46; Mismatches 69; Indels 13; Gaps 5;

QY 1 MTQILVDEONLARFLELELTHENVVDTEYDQDGLDKALSHYDYLIIIDMLPSNG 60
DB 1 MKKILVDEPAITLLQYNLEASHYQVETATDQGEALDKVRSEPFDFIILDLMLPSLSG 60

QY 61 LEICRKIRQOOSTPIIIITAKSDTYKVAGLDYGADDYIVKPFDEELLARIRAILR-- 118
DB 61 LDVTCKIREKIQTFPIMILTAQDNETKIVGLELGADDYVTKPFSPREIIARIKAIRRS 120

QY 119 --QPQKDI-----IDVNGITIDKNAFKVTVNGAEIELTKTEYDLYLLAENKHHVMORE 170
DB 121 QSQPQAPLSRNNAOITVGOITIDPENYKASKAGHRLQITPKFEFLLVYFAQRVGKTLSPD 180

QY 171 QILNHVWGYNSEVETNVVDVYIRYLNRKLPK--PYDRDKMIETVRGVY 216
DB 181 ALLNGVMGFDYPAETRMVDIQVSHLRDKIETDPKHPD--YLKTVRGFGY 227

RESULT 6
US-09-107-532A-4135
; Sequence 4135, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; CORRESPONDENCE ADDRESS:
; ADDRESS: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 4135:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 246 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
```



```
; ORGANISM: Enterococcus faecium
;
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...246
; SEQUENCE DESCRIPTION: SEQ ID NO: 4135:
US-09-107-532A-4135

Query Match      40.8%; Score 458; DB 4; Length 246;
Best Local Similarity 41.5%; Pred. No. 1.5e-40;
Matches 95; Conservative 47; Mismatches 77; Indels 10; Gaps 3;

Qy 1 MTQILIVEDEQNARLELELTHENVNVDTEYDQGLDKALSHYVDLIILDLMLPSING 60
Db 13 MKKVLVVDDKPSIDVKNLAKGYDYTAIDGEEALEKVAEVPDLILLDLMLPKMDG 72

Qy 61 LEICRKIROQOSTPIIIITAKSDTYDKVAGLDYGADDYIVKPPDIEELLARAILRR-- 118
Db 73 LEVAREVRKTYDMPILMVTAKDSIDKVLGLGELGADDYVTKPFSNRELVARVKANLRGA 132

Qy 119 -----QPKDI---IDVNGITIDKNAFKVTYNGAEIELTKTEYDLYLLAENKNHVMOREQ 171
Db 133 TAKEPEEAPAEUTIGDLTIHPEAYMVTGRGETIELTHREFELLYLAKHLGQVMTREH 192

Qy 172 ILNHWGYNSEVETNVVDVYIRLYRNKLKPY-DRDKRMETVRGCVGYVIR 219
Db 193 LLQTVMGYDFGDRVTVTVTRRLREKIEDNPSHPNVLVTRRGVGYLR 241

RESULT 7
US-09-134-001C-3175
; Sequence 3175, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 3175
; LENGTH: 246
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3175
```

```
Query Match      40.3%; Score 452; DB 4; Length 246;
Best Local Similarity 40.7%; Pred. No. 6.5e-40;
Matches 92; Conservative 49; Mismatches 73; Indels 12; Gaps 3;

Qy 3 QILIVEDEQNARLELELTHENVNVDTEYDQGLDKALSHYVDLIILDLMLPSINGLE 62
Db 9 EILIVDDERIRLLKLYLRESEPEIHEARDGKEAYELALENNYACILLDLMLPEMDGIE 68

Qy 63 ICRKIROQOSTPIIIITAKSDTYDKVAGLDYGADDYIVKPPDIEELLARAILRR--- 118
Db 69 VASKLREHKDTPILMTAKGEETNRVEGFEAGADDYIVKPPSPREVVLVRKALLRRQTQA 128

Qy 119 -----QPKDI---IDVNGITIDKNAFKVTYNGAEIELTKTEYDLYLLAENKNHVMOREQ 172
Db 129 NTEQSEPHARDIIEFSLVINDNDAHRLVADDQVNLTPKYEYELLYLAKTPNKFVREQ 188

Qy 173 ILNHWGYNSEVETNVVDVYIRLYRNKLKPYDRD--KMIEIVRGVGY 216
Db 189 LKEVWHYEFYDGLTVDTHVRLREKLNRRVSSDAQMIIQIVWGVGY 234
```

RESULT 8
US-09-107-532A-6746

```
; Sequence 6746, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneka
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 6746:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 234 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...234
; SEQUENCE DESCRIPTION: SEQ ID NO: 6746:
US-09-107-532A-6746
```

```
Query Match      40.1%; Score 449.5; DB 4; Length 234;
Best Local Similarity 39.6%; Pred. No. 1.1e-39;
Matches 91; Conservative 58; Mismatches 64; Indels 17; Gaps 5;

Qy 1 MTQILIVEDEQNARLELELTHENVNVDTEYDQGLDKALSHYVDLIILDLMLPSING 60
Db 1 MKKILVVDDPSIVTLTTFNLEKEGYKVTSAADGGEGLEALELQSPDFILDLVMTMDG 60

Qy 61 LEICRKIROQOS-TPIIITAKSDTYDKVAGLDYGADDYIVKPPDIEELLARAILRR- 118
Db 61 MEITQRLQEKNETPILMLTAKDQVDRIIGLEIGADDYLTTPKPSPREVLARKAIFRI 120

Qy 119 QPK-----DIIDVNGITIDKNAFKVTYNGAEIELTKTEYDLYLLAENKNHVMORE 170
Db 121 EPRKQSKDAEPYLSIGQITADLTNYQVTVDDHLLIETPKFELLVYFMKRDVRIDR 180

Qy 171 QILNHWGYNSEVETNVVDVYIRLYRNKLKPYDRD----KMIEIVRGVGY 216
Db 181 TLLDRWNFDFAQSQSRIVDVHSHLREKI---ERDPKPKYLVTVRGFGY 227
```

RESULT 9
US-09-082-077-4
; Sequence 4, Application US/09082077

Patent No. 6514746
; GENERAL INFORMATION:
; APPLICANT: Microcide Pharmaceuticals, Inc.
; TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS HISTIDINE PROTEIN KINASE
; FILE REFERENCE: SR Silverstein Microcide 234/06705
; CURRENT APPLICATION NUMBER: US/09/082,077
; PRIOR FILING DATE: 1998-05-20
; PRIOR APPLICATION NUMBER: US 60/009,102
; PRIOR FILING DATE: 1995-12-22
; PRIOR APPLICATION NUMBER: PCT/US97/23912
; PRIOR FILING DATE: 1997-12-23
; PRIOR APPLICATION NUMBER: US 08/713,718
; PRIOR FILING DATE: 1996-09-13
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 233
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-082-077-4

Query Match 40.0%; Score 449; DB 4; Length 233;
Best Local Similarity 40.5%; Pred. No. 1.2e-39;
Matches 92; Conservative 47; Mismatches 78; Indels 10; Gaps 3;

Qy 3 QILIVEDQNLFLELELTHENVNVDTEYDQGLDKALSHYDYLIIIDMLPDSINGLE 62
Db 4 KVVVDDEKPIADILEFNLKKEGYDVYCAVDGDAVDLIYEEEPDIVLLDMLPDRGME 63

Qy 63 ICRKIROQOSTPIIIITAKSDTYDKVAGLDYGADDYIVKPFDBIELLARIILRR---Q 119
Db 64 VCREVRKKYEMPIIMLTAKDSEIDKVLGLGAGDDYVTKPSTRELIARVKANLRRHSQ 123

Qy 120 PQKDI-----IDVNGITIDKNAKVTGNAEIELTKTEYDILLLAENKHNHVMQREOIL 173
Db 124 PAQDTGNVTNEITIKDVIYDPDAYSIKKRGEDIETHREFELFHYLSKHMGQVMTREHLL 183

Qy 174 NHVGYNSEVTNVVDVYIRYLRNKLK-PYDRDKMIETVRGVGVIR 219
Db 184 QTVWGYDFGDRVTDVTRIRLEKIEDDPSPHEIYIVTRGVGVFLQ 230

RESULT 10
US-09-134-001C-3063
; Sequence 3063, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 3063
; LENGTH: 234
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3063

Query Match 40.0%; Score 449; DB 4; Length 234;
Best Local Similarity 40.1%; Pred. No. 1.3e-39;
Matches 91; Conservative 49; Mismatches 77; Indels 10; Gaps 3;

Qy 3 QILIVEDQNLFLELELTHENVNVDTEYDQGLDKALSHYDYLIIIDMLPDSINGLE 62
Db 5 KVVVDDEKPIADILEFNLKKEGYDVYCAVDGDAVDLIYEEEPDIVLLDMLPDRGME 64

Qy 63 ICRKIROQOSTPIIIITAKSDTYDKVAGLDYGADDYIVKPFDBIELLARIILRR---Q 119
Db 65 VCREVRKKYEMPIIMLTAKDSEIDKVLGLGAGDDYVTKPSTRELIARVKANLRRHSQ 124

Qy 120 PQKDI-----IDVNGITIDKNAKVTGNAEIELTKTEYDILLLAENKHNHVMQREOIL 173
Db 125 PAQSVGATNEITIKDVIYDPDAYSIKKRGEDIETHREFELFHYLSKHMGQVMTREHLL 184

Qy 174 NHVGYNSEVTNVVDVYIRYLRNKLK-PYDRDKMIETVRGVGVIR 219
Db 185 QTVWGYDFGDRVTDVTRIRLEKIEDDPSPHEIYIVTRGVGVFLQ 231

RESULT 11
US-08-956-171E-5239
; Sequence 5239, Application US/08956171E
; Patent No. 6593114
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; Gil H. Choi
; Patrick S. Dillon
; Craig A. Rosen
; Steven C. Barash
; Michael R. Fannon
; ADDRESS: Human Genome Sciences, Inc.
; ADDRESS: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/956,171E
; FILING DATE: 20-Oct-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/009,861
; FILING DATE: January 5, 1996
; APPLICATION NUMBER: 08/781,986
; FILING DATE: January 3, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mark J. Hyman
; REGISTRATION NUMBER: 46,789
; REFERENCE/DOCKET NUMBER: PB248P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (240) 314-1224
; TELEFAX: (301) 309-8439
; INFORMATION FOR SEQ ID NO: 5239:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 239 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 5239:
US-08-956-171E-5239

Query Match 40.0%; Score 449; DB 4; Length 239;
Best Local Similarity 40.5%; Pred. No. 1.3e-39;
Matches 92; Conservative 47; Mismatches 78; Indels 10; Gaps 3;

Qy 3 QILIVEDQNLFLELELTHENVNVDTEYDQGLDKALSHYDYLIIIDMLPDSINGLE 62
Db 10 KVVVDDEKPIADILEFNLKKEGYDVYCAVDGDAVDLIYEEEPDIVLLDMLPDRGME 69

Qy 63 ICRKIROQOSTPIIIITAKSDTYDKVAGLDYGADDYIVKPFDBIELLARIILRR---Q 119

Db 70 VCREVRKYEPIIMLTAKOSEIDKVLGLGAGDDYVTKPFSRELLIARVKNLRHYSQ 129
Qy 120 PQKDI-----IDVNGTIDKNAPKVTNGAGIELTKTEYDLYLLAENKNHVMQREIIL 173
Db 130 PAQDTGNVTNEITIKDIVIPDAYSIKRGEDIELTHREFELFHYLSKMGQVMTREHLL 189
Qy 174 NHWGYNSEVETNVVDVYIRYLRNKLK-PYDRDKMIETVRGVGVIR 219
Db 190 QTVNGYDFGVRTVDVTVRRLEKIEDPSHEIYVTRRGVGFLLQ 236
RESULT 12
US-08-850-118-2
; Sequence 2, Application US/08850118
; Patent No. 5773250
; GENERAL INFORMATION:
; APPLICANT: Wallis, Nicola
; TITLE OF INVENTION: No. 5773250el Compounds
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/850,118
; FILING DATE: 01-MAY-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 9609021.2
; FILING DATE: 01-MAY-1996
; NAME: Gimmi, Edward R
; ATTORNEY/AGENT INFORMATION:
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: P31492
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 234 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-850-118-2
Query Match 39.4%; Score 442; DB 1; Length 234;
Best Local Similarity 40.3%; Pred. No. 6.9e-39;
Matches 94; Conservative 49; Mismatches 74; Indels 16; Gaps 3;
Qy 1 MTQILIVEDONLARFLELELTHENVNVDTEYDGDGLDKALSHYVDLIILDLMLPSING 60
Db 1 MIKILIVDDKPSIDIIKFNMTKEGYEVVTAFGREALEQFEAEQPDIIILDLMLPEIDG 60
Qy 61 LEICRKIROQOSTPIIIITAKSDTYDKVAGLDYGDADYIVKPFDIIEELLARAILRR-- 118
Db 61 LEVAKTIRKTSVPFILMSAKDSEFDKVLGLEIGADDYVTKPFSNRELQARVKALLRRSQ 120
Qy 119 -----OPQKDIIDVNGITIDKNAPKVTNGAGIELTKTEYDLYLLAENKNHVM 167
Db 121 PMPVDGQEADSKPQP--IQIGDLIEVPDAYVAKYKGEELDLTHREFELLYHLASHTGQVI 178
Qy 168 QREQILNHWGYNSEVETNVVDVYIRYLRNKLKPY-DRDKMIETVRGVGVIR 219
Db 190 QTVNGYDFGVRTVDVTVRRLEKIEDPSHEIYVTRRGVGFLLQ 236
RESULT 13
US-09-008-253-2
; Sequence 2, Application US/09008253
; Patent No. 5910572
; GENERAL INFORMATION:
; APPLICANT: Wallis, Nicola
; TITLE OF INVENTION: No. 5910572el Compounds
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/008,253
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/850,118
; FILING DATE: 01-MAY-1997
; APPLICATION NUMBER: 9609021.2
; FILING DATE: 01-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Gimmi, Edward R
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: P31492
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 234 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-008-253-2
Query Match 39.4%; Score 442; DB 2; Length 234;
Best Local Similarity 40.3%; Pred. No. 6.9e-39;
Matches 94; Conservative 49; Mismatches 74; Indels 16; Gaps 3;
Qy 1 MTQILIVEDONLARFLELELTHENVNVDTEYDGDGLDKALSHYVDLIILDLMLPSING 60
Db 1 MIKILIVDDKPSIDIIKFNMTKEGYEVVTAFGREALEQFEAEQPDIIILDLMLPEIDG 60
Qy 61 LEICRKIROQOSTPIIIITAKSDTYDKVAGLDYGDADYIVKPFDIIEELLARAILRR-- 118
Db 61 LEVAKTIRKTSVPFILMSAKDSEFDKVLGLEIGADDYVTKPFSNRELQARVKALLRRSQ 120
Qy 119 -----OPQKDIIDVNGITIDKNAPKVTNGAGIELTKTEYDLYLLAENKNHVM 167
Db 121 PMPVDGQEADSKPQP--IQIGDLIEVPDAYVAKYKGEELDLTHREFELLYHLASHTGQVI 178
Qy 168 QREQILNHWGYNSEVETNVVDVYIRYLRNKLKPY-DRDKMIETVRGVGVIR 219
Db 190 QTVNGYDFGVRTVDVTVRRLEKIEDPSHEIYVTRRGVGFLLQ 236
RESULT 14
US-09-093-335-2

Db 179 TREHLLSTWGVYDFGVRTVDVTVRRLEKIEDTSPREYILTRRGVGYMR 231
RESULT 13
US-09-008-253-2
; Sequence 2, Application US/09008253
; Patent No. 5910572
; GENERAL INFORMATION:
; APPLICANT: Wallis, Nicola
; TITLE OF INVENTION: No. 5910572el Compounds
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/008,253
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/850,118
; FILING DATE: 01-MAY-1997
; APPLICATION NUMBER: 9609021.2
; FILING DATE: 01-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Gimmi, Edward R
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: P31492
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 234 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-008-253-2
Query Match 39.4%; Score 442; DB 2; Length 234;
Best Local Similarity 40.3%; Pred. No. 6.9e-39;
Matches 94; Conservative 49; Mismatches 74; Indels 16; Gaps 3;
Qy 1 MTQILIVEDONLARFLELELTHENVNVDTEYDGDGLDKALSHYVDLIILDLMLPSING 60
Db 1 MIKILIVDDKPSIDIIKFNMTKEGYEVVTAFGREALEQFEAEQPDIIILDLMLPEIDG 60
Qy 61 LEICRKIROQOSTPIIIITAKSDTYDKVAGLDYGDADYIVKPFDIIEELLARAILRR-- 118
Db 61 LEVAKTIRKTSVPFILMSAKDSEFDKVLGLEIGADDYVTKPFSNRELQARVKALLRRSQ 120
Qy 119 -----OPQKDIIDVNGITIDKNAPKVTNGAGIELTKTEYDLYLLAENKNHVM 167
Db 121 PMPVDGQEADSKPQP--IQIGDLIEVPDAYVAKYKGEELDLTHREFELLYHLASHTGQVI 178
Qy 168 QREQILNHWGYNSEVETNVVDVYIRYLRNKLKPY-DRDKMIETVRGVGVIR 219
Db 179 TREHLLSTWGVYDFGVRTVDVTVRRLEKIEDTSPREYILTRRGVGYMR 231
RESULT 14
US-09-093-335-2

Sequence 2, Application US/09093335
Patent No. 6001362
GENERAL INFORMATION:
APPLICANT: Wallis, Nicola
TITLE OF INVENTION: No. 6001362el Compounds
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: SmithKline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: USA
ZIP: 19406-0939
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA: US/09/093.335
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/850.118
FILING DATE: 01-MAY-1997
APPLICATION NUMBER: 9609021.2
FILING DATE: 01-MAY-1996
ATTORNEY/AGENT INFORMATION:
NAME: Gimmi, Edward R.
REGISTRATION NUMBER: 38,891
REFERENCE/DOCKET NUMBER: P31492
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-4478
TELEFAX: 610-270-5090
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 234 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-093-335-2

Query Match 39.4%; Score 442; DB 3; Length 234;
Best Local Similarity 40.3%; Pred. No. 6.9e-39;
Matches 94; Conservative 49; Mismatches 74; Indels 16; Gaps 3;
Qy 1 MTQILIVDEQNARFLELELTHENVNVDYDQGLDKALSHYYDLIILDMLPSING 60
Db 1 MIKILIVDDEKPISDIIRKFNMTKEGYEVVTAFGREALEQFEAEQPDIIILDMLPEIDG 60
Qy 61 LEICRKIRQOOSTPIIIITAKSDTYDVGADLYGADYIVKPFDEIELLARAILRR-- 118
Db 61 LEVAKTIRKTSVPILMLSAKDSDFKVLGELGADYVTKPFNSRELQARVKALLRSQ 120
Qy 119 -----OPQKDIIDVNGITIDKNAFKVTNGAGIELTKTEYDLYLLAENKNHVM 167
Db 121 PMPVDGQADSKQP--IQIGDLVDPAYVAKYGEELDLTHREFELLYHLASHTGGVI 178
Qy 168 QREQLNHVWYNGSEVETNVVDVYIRYLNRNKLKPY-DRDKMIETVRGVGYVIR 219
Db 179 TREHLETVWGYDYGVDVTRVRLREKIEDTPSRPEVILTRGVGYVYMR 231

RESULT 15
US-09-134-001C-3779
Sequence 3779, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS

FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134.001C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 3779
LENGTH: 245
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3779

Query Match 39.4%; Score 442; DB 4; Length 245;
Best Local Similarity 40.3%; Pred. No. 7.4e-39;
Matches 91; Conservative 53; Mismatches 70; Indels 12; Gaps 4;
Qy 3 QILIVDEQNARFLELELTHENVNVDYDQGLDKALSHYYDLIILDMLPSINGLE 62
Db 13 KILVDDDEQSIVITLLKYNLETAGYIVVAYDGEALKKVETEPELIVLDVNLPKKDGIE 72
Qy 63 ICRKIRQOOS-TPIIITAKSDTYDVGADLYGADYIVKPFDEIELLARAILRR--- 118
Db 73 VCKTIRSDXNLVPIILMTAKDDDFRVLGELGADYVTKPFSPREVVARVKAILRRSQF 132
Qy 119 --QPQKDIID-----VNGITIDKNAFKVTNGAGIELTKTEYDLYLLAENKNHVMOREQ 171
Db 133 VNEIEKEDVDEDIIGSIRIRPEFFEVYKEDELELTPEFELLYLIERQGRVITREH 192
Qy 172 ILNHVWYNGSEVETNVVDVYIRYLNRNKLKPYDRD-KMIETVRGVGY 216
Db 193 MLNSVMWYEFAGDSRIVDVHSHLRDKLEENPKQPKLIKTVRGLGY 238

Search completed: October 4, 2004, 18:11:19
Job time : 18.0164 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 4, 2004, 18:09:40 ; Search time 58.509 Seconds
(without alignments)
1204.500 Million cell updates/sec

Title: US-09-006-627-4

Perfect score: 1122

Sequence: 1 MTQILVEDQNLFARLEFE.....KPYDRDRKMTVRGVGYR 219

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1351062 seqs, 321799191 residues

Total number of hits satisfying chosen parameters: 1351062

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.*

- 1: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pdb.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pdb.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pdb.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pdb.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pdb.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pdb.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pdb.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pdb.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pdb.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pdb.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pdb.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pdb.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pdb.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pdb.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pdb.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pdb.*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pdb.*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pdb.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	498	44.4	230	12	US-10-282-122A-53520
2	490.5	43.7	232	12	US-10-282-122A-51564
3	476	42.4	232	12	US-10-282-122A-52443
4	472	42.1	243	14	US-10-156-761-11581
5	468.5	41.8	234	15	US-10-264-213-128
6	461.5	41.1	226	9	US-09-990-337-2
7	461.5	41.1	226	9	US-09-738-626-4340
8	461.5	41.1	226	14	US-10-411-318-2
9	458	40.8	261	12	US-10-282-122A-57574
10	457	40.7	232	9	US-09-738-626-4476
11	456.5	40.7	235	12	US-10-282-122A-46382
12	456	40.6	234	9	US-09-815-242-10748
13	456	40.6	234	12	US-10-282-122A-42560
14	452	40.3	233	9	US-09-815-242-4913
15	449	40.0	233	9	US-09-815-242-5676

16	449	40.0	233	9	US-09-925-637-46
17	449	40.0	233	14	US-10-084-205-46
18	449	40.0	233	16	US-10-712-713-46
19	449	40.0	235	9	US-09-815-242-12674
20	449	40.0	235	9	US-09-815-242-13000
21	449	40.0	235	12	US-10-282-122A-44277
22	449	40.0	239	8	US-08-781-986A-5239
23	449	40.0	239	12	US-10-282-122A-70821
24	449	40.0	239	12	US-10-329-624-5239
25	447	39.8	238	12	US-10-282-122A-71628
26	443	39.5	245	14	US-10-156-761-12301
27	442.5	39.4	228	12	US-10-282-122A-63751
28	442	39.4	234	9	US-09-815-242-13285
29	442	39.4	234	9	US-09-815-242-13558
30	442	39.4	251	12	US-10-282-122A-74003
31	439.5	39.2	229	14	US-10-156-761-12597
32	436.5	38.9	235	12	US-10-282-122A-72172
33	435	38.8	223	14	US-10-156-761-11506
34	433	38.6	236	12	US-10-282-122A-74434
35	432.5	38.5	237	12	US-10-282-122A-60754
36	430.5	38.4	239	14	US-10-287-274-367
37	425.5	37.9	269	14	US-10-156-761-11732
38	424	37.8	226	12	US-10-282-122A-53860
39	421	37.5	238	12	US-10-282-122A-59662
40	419.5	37.4	225	10	US-09-882-227-486
41	419	37.3	272	12	US-10-282-122A-48229
42	417.5	37.2	225	9	US-09-815-242-11478
43	417.5	37.2	225	12	US-10-335-977-7469
44	417.5	37.2	226	12	US-10-335-977-7470
45	414	36.9	225	14	US-10-156-761-11018

ALIGNMENTS

RESULT 1

- US-10-282-122A-53520
- Sequence 53520, Application US/10282122A
- Publication No. US20040029129A1
- GENERAL INFORMATION:
- APPLICANT: Wang, Liangsu
- APPLICANT: Zamudio, Carlos
- APPLICANT: Malone, Cheryl
- APPLICANT: Haselbeck, Robert
- APPLICANT: Ohlsen, Kari
- APPLICANT: Zvekind, Judith
- APPLICANT: Wall, Daniel
- APPLICANT: Trawick, John
- APPLICANT: Carr, Grant
- APPLICANT: Yamamoto, Robert
- APPLICANT: Forsyth, R.
- APPLICANT: Xu, H
- TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
- FILE REFERENCE: ELITRA.034A
- CURRENT APPLICATION NUMBER: US/10/282,122A
- CURRENT FILING DATE: 2003-02-20
- PRIOR APPLICATION NUMBER: 60/191,078
- PRIOR FILING DATE: 2000-03-21
- PRIOR APPLICATION NUMBER: 60/206,848
- PRIOR FILING DATE: 2000-05-23
- PRIOR APPLICATION NUMBER: 60/207,727
- PRIOR FILING DATE: 2000-05-26
- PRIOR APPLICATION NUMBER: 60/230,335
- PRIOR FILING DATE: 2000-09-06
- PRIOR APPLICATION NUMBER: 60/230,347
- PRIOR FILING DATE: 2000-09-09
- PRIOR APPLICATION NUMBER: 60/242,578
- PRIOR FILING DATE: 2000-10-23
- PRIOR APPLICATION NUMBER: 60/253,625
- PRIOR FILING DATE: 2000-11-27
- PRIOR APPLICATION NUMBER: 60/257,931
- PRIOR FILING DATE: 2000-12-22
- PRIOR APPLICATION NUMBER: 60/267,636


```
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 52443
; LENGTH: 232
; TYPE: PRP
; ORGANISM: Clostridium botulinum
US-10-282-122A-52443

Query Match          42.4%; Score 476; DB 12; Length 232;
Best Local Similarity 43.3%; Pred. No. 6.6e-40;
Matches 97; Conservative 51; Mismatches 68; Indels 8; Gaps 3;

Qy 1 MTQILIVEDEQNLFARFLELTHENVNVDTEYDQDGLDKALSHYDYLIDMLPSSING 60
Db 3 MEKILIVDEDEHICELIKFNLENNSYNCICALNGIEALNIAKEKPDILIDMLPGMDG 62

Qy 61 LEICRKIRQQQ---STPIIIITAKSDTYDKVAGLDYGDADDYIVKPFDEIELLARILRR 117
Db 63 YEVCKEIRKONNIATTPIMITAKGEELDKVGLGELGADDYITKPFVSRENVARVAVLR 122

Qy 118 RQPKQDIID----VNGITIDKNAPKVTNVGAEIELTKTEYDLYLLAENKHNVMQREQL 173
Db 123 RRTINIESKSYILGNIVDPEKHSITKAGEKIELTKPELLEVLKNGKGVNTRDFLL 182

Qy 174 LNHVWGYNSEVTNVVDVYIRYLRNKLKPYDRD-KMIETVRGVGY 216
Db 183 DRWNGVEYIGETRTVDVHRLRKKIEDDDDKNPKFIETIRGIGY 226

RESULT 4
US-10-156-761-11581
; Sequence 11581, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 11581
; LENGTH: 243
; TYPE: PRP
; ORGANISM: Streptomyces avermitilis
US-10-156-761-11581

Query Match          42.1%; Score 472; DB 14; Length 243;
Best Local Similarity 41.4%; Pred. No. 1.8e-39;
Matches 94; Conservative 48; Mismatches 75; Indels 10; Gaps 2;

Qy 3 QILIVEDEQNLFARFLELTHENVNVDTEYDQDGLDKALSHYDYLIDMLPSSINGLE 62
Db 12 RILIVDEPAVREALQSLAFEGYDTEVAVDGALEKATAYQDVLVLDIQMPMDGLT 71

Qy 63 ICRKIR-QQOSTPIIIITAKSDTYDKVAGLDYGDADDYIVKPFDEIELLARILRRQ-- 119
Db 72 AARRIRGAGTTPTIMLTARDTVGDRVTGLDAGADDYLVKPFDELDFARILLRRSSY 131

Qy 120 -----PQKDIIDVNGITIDKNAPKVTNVGAEIELTKTEYDLYLLAENKHNVMQREQL 172
Db 132 AATAVGPEDDALTFGDLRMDLATREVRGRSSVELIRTEFTLLEMFALHPQVLTREQI 191

Qy 173 LNHVWGYNSEVTNVVDVYIRYLRNKLKPYDRDKMIETVRGVGYVIR 219
```

```
Db 192 LKAVWGFDPFPSSNSLDVYVYMLRRKTEAGGPELRLVHTVRGVGYVLR 238

RESULT 5
US-10-264-213-128
; Sequence 128, Application US/10264213
; Publication No. US20040009490A1
; GENERAL INFORMATION:
; APPLICANT: Glenn, Matthew
; APPLICANT: Havukkala, Ilkka J
; APPLICANT: Lubbers, Mark William
; APPLICANT: Dekker, James
; TITLE OF INVENTION: Polynucleotides, materials incorporating
; TITLE OF INVENTION: them, and methods for using them.
; FILE REFERENCE: 11000.1043c3
; CURRENT APPLICATION NUMBER: US/10/264,213
; CURRENT FILING DATE: 2002-10-03
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 128
; LENGTH: 234
; TYPE: PRP
; ORGANISM: Lactobacillus rhamnosus
US-10-264-213-128

Query Match          41.8%; Score 468.5; DB 15; Length 234;
Best Local Similarity 43.9%; Pred. No. 3.9e-39;
Matches 100; Conservative 46; Mismatches 69; Indels 13; Gaps 5;

Qy 1 MTQILIVEDEQNLFARFLELTHENVNVDTEYDQDGLDKALSHYDYLIDMLPSSING 60
Db 1 MKXILIVDDPEAITLTLLQYNLEAEHYQVETATDQEALDKVRSPPDFIILDLMLPSLSG 60

Qy 61 LEICRKIRQQQ-STPIIIITAKSDTYDKVAGLDYGDADDYIVKPFDEIELLARILRR- 118
Db 61 LDVTKIREKIQTPIMLTAKDNETDKIVGLGELGADDYVTKPFSPEIIRIKAIERRS 120

Qy 119 --OPQKDI-----IDVNGITIDKNAPKVTNVGAEIELTKTEYDLYLLAENKHNVMQRE 170
Db 121 QSQPQAPLSRANQITVGQITIDPENYKASKAGHRLQUTPKFELLVYFQRVGKTLSDR 180

Qy 171 QILNHVWGYNSEVTNVVDVYIRYLRNKLK--PYDRDKMIETVRGVGY 216
Db 181 ALLNGVWGFDPYPAETRMVDIQVSHLRDKIETDPKHPD-VLKTVRGVGY 227

RESULT 6
US-09-990-337-2
; Sequence 2, Application US/09990337
; Patent No. US20020137073A1
; GENERAL INFORMATION:
; APPLICANT: BATHE, Brigitte
; APPLICANT: SCHISCHKA, Natalie
; APPLICANT: BOTT, Michael
; APPLICANT: SCHAPPER, Steffan
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES CODING FOR THE MtrA and/or MtrB PROTEINS
; FILE REFERENCE: 215474USOX
; CURRENT APPLICATION NUMBER: US/09/990,337
; CURRENT FILING DATE: 2001-11-23
; PRIOR APPLICATION NUMBER: DE10057802.0
; PRIOR FILING DATE: 2000-11-22
; PRIOR APPLICATION NUMBER: DE10125089.4
; PRIOR FILING DATE: 2001-05-23
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 226
; TYPE: PRP
; ORGANISM: Corynebacterium glutamicum
US-09-990-337-2

Query Match          41.1%; Score 461.5; DB 9; Length 226;
```



```
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 57574
; LENGTH: 261
; TYPE: PRT
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (18)..(18)
; OTHER INFORMATION: X=any amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (20)..(20)
; OTHER INFORMATION: X=any amino acid
; US-10-282-122A-57574

Query Match          40.8%; Score 458; DB 12; Length 261;
Best Local Similarity 41.5%; Pred. No. 5.3e-38;
Matches 95; Conservative 47; Mismatches 77; Indels 10; Gaps 3;

Qy 1 MTQILVEDEQNARFLLELTHENYNVDTEYDQGLDKALSHYDILILDLMLPSING 60
Db 28 MKVLVWDEKPISDIVKFNLAKEGYDVTAYDGEAELEKVAEVEPDILILDLMLPKWDG 87
Qy 61 LEICRKIROQOSTPIIIITAKSDTYDKVAGLDYDADDYIVKPFDEIELLARIRAILRR-- 118
Db 88 LEVAREVRKTYDPIIMVYTAQSDSIDKVLGLGADDDYVTKPFSNRELVARVKANLRGA 147
Qy 119 ----OPOKDI---IDVNGITIDKNKAFKVTVNGAEITLTKTEYDILLYLAENKHNWOREQ 171
Db 148 TAAKEPEEAAAEITIGDLTHPEAYVTKKGETIELTHREFELLYLAKHLGQVMTREH 207
Qy 172 ILNHVWGYNSVETNVVDVYIRYLRNKLKPY-DRDKMIETVRGVGVYIR 219
Db 208 LLQTVWGYDYFDVTRVTVTRRLREKIEDNPSPNPNVLTVRGVGVYLR 256

RESULT 10
US-09-738-626-4476
; Sequence 4476, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738.626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
```

```
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: Patent in ver. 3.0
; SEQ ID NO 4476
; LENGTH: 232
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
; US-09-738-626-4476

Query Match          40.7%; Score 457; DB 9; Length 232;
Best Local Similarity 41.9%; Pred. No. 5.6e-38;
Matches 95; Conservative 44; Mismatches 78; Indels 10; Gaps 2;

Qy 3 QILIVEDEQNARFLLELTHENYNVDTEYDQGLDKALSHYDILILDLMLPSINGLE 62
Db 2 KILVVDDEQAVRDSLRSLSFNGYVVLAEEDGTALEMDKEQPALVILDMVFMFGMDGLE 61
Qy 63 ICRKIROQ-QSTPIIIITAKSDTYDKVAGLDYDADDYIVKPFDEIELLARIRAILRRQP- 120
Db 62 VCHLRSEGDDRPILILITARDNVSDRVGGDAGADDYLAKPFALLEELARVSLVRSAV 121
Qy 121 -----QKDIIDVNGITIDKNKAFKVTVNGAEITLTKTEYDILLYLAENKHNWOREQI 172
Db 122 ESNQSSIEQALLSCGDLTLDPSRDVYRNGRAISLRTFTEFALLQLLKQKQKVLTRAQI 181
Qy 173 ILNHVWGYNSVETNVVDVYIRYLRNKLKPYDRDKMIETVRGVGVYIR 219
Db 182 LEEVWGCDFPTSGNALEVYIGYLRKTELEGEDRLIHTVRGVGVYLR 228

RESULT 11
US-10-282-122A-46382
; Sequence 46382, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282.122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
```



```
Qy 1 MTQILIVEQNLARFLELELTHENVNVDTEYDQGLDKALSHYVDLIILDLMLPSING 60
Db 1 MKKILVDDKPISEIVKYNLKEGEVFTAYDGEALEKVEVEPDLIILDLMLPKWDG 60

Qy 61 LEICRKIROQOSTPIIIITAKSDTYDKVAGLDYGADDYIVKPFDBELLARIRAILRR-- 118
Db 61 LEVAREVTKTHDPIIMVTAKSDSIDKVLGELGADDYVTKPFNSRELVARVKANLRGA 120

Qy 119 -----QPQKDIIDVNGITIDKNAKVTNVGAELTKTEYDLYLLAENKHNVMOREQ 171
Db 121 TNAKEAVTTQSELTIGDLTIHPDAYWVSKGEKIELTHREFELLYLAKHIGQVMTREH 180

Qy 172 ILNHVWGNSEVETNVVDVYIRYLRNKLKPY-DRDKMIETVRGVGVIR 219
Db 181 LQTVWGYDFGVRTVDVTVRRLREKIEDSPSHPTYLTVTRRGVGYLR 229

RESULT 14
US-09-815-242-4913
; Sequence 4913, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4913
; LENGTH: 233
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-815-242-4913

Query Match 40.3%; Score 452; DB 9; Length 233;
Best Local Similarity 41.9%; Pred. No. 1.8e-37;
Matches 95; Conservative 44; Mismatches 78; Indels 10; Gaps 2;

Qy 3 QILVVEQNLARFLELELTHENVNVDTEYDQGLDKALSHYVDLIILDLMLPSINGLE 62
Db 2 KILVVDDEKPISEIVKYNLKEGEVFTAYDGEALEKVEVEPDLIILDLMLPKWDGLE 61

Qy 63 ICRKIROQOSTPIIIITAKSDTYDKVAGLDYGADDYIVKPFDBELLARIRAILRR--- 118
Db 62 VAREVRKTHDPIIMVTAKSDSIDKVLGELGADDYVTKPFNSRELVARVKANLRGATN 121

Qy 119 -----QPQKDIIDVNGITIDKNAKVTNVGAELTKTEYDLYLLAENKHNVMOREQ 173
Db 122 AKEAEVTTQSELTIGDLTIHPDAYWVSKRGEKIELTHREFELLYLAKHIGQVMTREHLL 181
```

```
Qy 174 NHWGNVSEVETNVVDVYIRYLRNKLKPY-DRDKMIETVRGVGVIR 219
Db 182 QTVWGYDFGVRTVDVTVRRLREKIEDSPSHPTYLTVTRRGVGYLR 228

RESULT 15
US-09-815-242-5676
; Sequence 5676, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5676
; LENGTH: 233
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-5676

Query Match 40.0%; Score 449; DB 9; Length 233;
Best Local Similarity 40.5%; Pred. No. 3.7e-37;
Matches 92; Conservative 47; Mismatches 78; Indels 10; Gaps 3;

Qy 3 QILVVEQNLARFLELELTHENVNVDTEYDQGLDKALSHYVDLIILDLMLPSINGLE 62
Db 4 KVVVDDEKPIADIILEFNLKKEGYDVYCAVDGNDVLDIYEEEPDILVLLDMLPGRGME 63

Qy 63 ICRKIROQOSTPIIIITAKSDTYDKVAGLDYGADDYIVKPFDBELLARIRAILRR---Q 119
Db 64 VCREVRKKYEMPIIMLTAKDSEIDKVLGELGADDYVTKPFSTRELIARVKANLRHYSQ 123

Qy 120 PQKDI-----IDVNGITIDKNAKVTNVGAELTKTEYDLYLLAENKHNVMOREQ 173
Db 124 PAQDTGNVTNEITIKDIVIYPDAYSIKKRGEDIELTHREFELPHYLSKHMGMQVMTREHLL 183

Qy 174 NHWGNVSEVETNVVDVYIRYLRNKLK- PYDRDKMIETVRGVGVIR 219
Db 184 QTVWGYDFGVRTVDVTVRRLREKIEDSPSHPTYLTVTRRGVGYFLQ 230

Search completed: October 4, 2004, 18:26:28
Job time : 61.509 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 5, 2004, 00:12:59 ; Search time 301.968 Seconds
(without alignments)
10354.322 Million cell updates/sec

Title: US-09-006-627-3
Perfect score: 736
Sequence: 1 ATTACGTTTGTCACTGA.....TTAAATTTGATTTTCTA 736

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 337863 seqs, 212409041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_29Jan04:*

- 1: Geneseqn1980s:*
- 2: Geneseqn1990s:*
- 3: Geneseqn2000s:*
- 4: Geneseqn2001as:*
- 5: Geneseqn2001bs:*
- 6: Geneseqn2002as:*
- 7: Geneseqn2003as:*
- 8: Geneseqn2003bs:*
- 9: Geneseqn2003cs:*
- 10: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	736	100.0	736	2	AAT72328 Staphyloc
2	736	100.0	736	2	AAX90602 Cognate r
3	736	100.0	2201	2	AAX90601 Histidine
4	726	98.6	1440	2	AAT75036 Staphyloc
5	657	89.3	657	7	ACF74591 Staphyloc
6	406.4	55.2	3099	4	AH54422 S. epider
7	401.8	54.6	669	6	ABN90912 Staphyloc
8	390.4	53.0	3230	4	AH54308 S. epider
9	347	47.1	3018	4	AH54320 S. epider
10	302.6	41.1	501	4	AH53697 S. epider
11	302.6	41.1	501	4	AH53015 S. epider
12	224.8	30.5	110000	6	ABQ69245 14
13	224.8	30.5	110000	6	ABQ67195 3
14	202.2	27.5	110000	6	ABA03041 13
15	202.2	27.5	110000	6	ABN03041 14
16	193.6	26.3	110000	6	ABN71527 15
17	193.6	26.3	110000	6	ABN90521 16
18	193	26.2	687	6	ABN68466
19	193	26.2	1908	2	AAX12964 Enterococ
20	193	26.2	1908	6	ABN98759 Enterococ
21	187.2	25.4	687	9	AUC93255 E. faeciu
22	184.2	25.0	110000	6	ABQ69245 13
23	173.8	23.6	684	6	ABN67673 Streptoco

C 24	165.6	22.5	110000	8	ADB12064_08	Continuation (9 of
C 25	161	21.9	705	8	ADB09375	Abn90875 Staphyloc
C 26	149.8	20.4	741	6	ABN90875	Abn90875 Staphyloc
C 27	146.6	19.9	672	6	ABN67082	Abn67082 Streptoco
C 28	144	18.6	684	8	ADA32762	ADA32762 DNA encod
C 29	138.8	18.9	110000	6	ABA90521_17	Continuation (18 o
C 30	137.6	18.7	110000	6	ABQ69245_20	Continuation (21 o
C 31	137.2	18.6	110000	6	ABN71527_20	Continuation (21 o
C 32	135.6	18.4	741	7	ACF72642	ACF72642 Staphyloc
C 33	131.8	17.9	690	7	ACA29466	ACA29466 Prokaryot
C 34	131	17.8	110000	6	ABN71527_06	Continuation (7 of
C 35	130.2	17.7	708	6	ABN66633	ABN66633 Streptoco
C 36	130	17.7	31517	2	AAX13117	AAX13117 Enterococ
C 37	130	17.7	31517	6	ABN98912	ABN98912 Enterococ
C 38	128	17.4	1037	6	ABQ70350	Abq70350 Listeria
C 39	127.4	17.3	900	2	AAT72329	AAT72329 Staphyloc
C 40	126.4	17.2	717	6	ABQ68073	ABQ68073 Listeria
C 41	126.4	17.2	110000	6	ABN71527_09	Continuation (10 o
C 42	126.4	17.2	110000	6	ABA03041_10	Continuation (11 o
C 43	126.4	17.2	110000	6	ABA03041_20	Continuation (21 o
C 44	125.8	17.1	12022	2	AAX13054	Aax13054 Enterococ
C 45	125.8	17.1	12022	6	ABS98849	ABS98849 Enterococ

ALIGNMENTS

RESULT 1	
AAT72328	
ID	AAT72328 standard; cDNA; 736 BP.
XX	
AC	AAT72328;
XX	
DT	17-OCT-2003 (revised)
DT	16-JAN-1998 (first entry)
XX	
DE	Staphylococcus aureus cDNA encoding a novel response regulator protein.
XX	
KW	Antibacterial; anti-Staphylococcus; gene therapy; bacterial infection;
KW	prophylaxis; two component signal transduction system; TCSTS; Read;
KW	response regulator; Staphylococcus; Bacillus; vaccine; wound treatment;
KW	dental; surgery; antibody; ds.
XX	
OS	Staphylococcus aureus; WCUH29.
XX	
FH	Key
FT	Location/Qualifiers
FT	Complement(13. .672)
FT	/*tag= a
FT	/product= "response_regulator_protein"
XX	
PN	WO9723506-A1.
XX	
PD	03-JUL-1997.
XX	
PP	20-DEC-1996; 96WO-GB003261.
XX	
PR	22-DEC-1995; 95GB-00026359.
XX	
PA	(SMIK) SMITHKLINE BEECHAM PLC.
XX	
PI	Wallis N, Hodgson JE;
XX	
DR	WPI; 1997-350970/32.
XX	
DR	P-PSDB; AAW19274.
XX	
PT	Isolated nucleic acid encoding Staphylococcus aureus response regulator -
PT	useful to treat or prevent bacterial infections, in vaccines and for
PT	diagnosis.
XX	
PS	Claim 4; Page 45-46; 59pp; English.
XX	
CC	The present sequence represents a polynucleotide, encoding a novel
CC	response regulator protein from Staphylococcus aureus. DNA was isolated


```
Qy 61 ACGATCGTATGTTTAACTTTGTTTCGTAATATCTTATATAAATCTACGACATTTGT 120
Db 61 ACGATCGTATGTTTAACTTTGTTTCGTAATATCTTATATAAATCTACGACATTTGT 120
Qy 121 TTCTACTTCATATATATACCCCATACATGATTTAAATTTGTTCCCGTTGCATAACATG 180
Db 121 TTCTACTTCATATATATACCCCATACATGATTTAAATTTGTTCCCGTTGCATAACATG 180
Qy 181 GTTTTATTTTTCAGCTAGATATAGTAAATCATACCTCTGTTTGTGTAATCAATTC 240
Db 181 GTTTTATTTTTCAGCTAGATATAGTAAATCATACCTCTGTTTGTGTAATCAATTC 240
Qy 241 TGGCCATTTACCGTCACCTTTAAAGCGTTCTTATCAATTTGAATACCGTTGACATCAT 300
Db 241 TGGCCATTTACCGTCACCTTTAAAGCGTTCTTATCAATTTGAATACCGTTGACATCAT 300
Qy 301 AATATCCTTTTGTGGCTGACGACGTAATTTGACGAAATTTCTGTAAGGTTCTTCAAT 360
Db 301 AATATCCTTTTGTGGCTGACGACGTAATTTGACGAAATTTCTGTAAGGTTCTTCAAT 360
Qy 361 ATCAAAACGGCTTAACATATATATCTGTCGACCGTAATCAAGCCGACCACTTTGTCTATA 420
Db 361 ATCAAAACGGCTTAACATATATATCTGTCGACCGTAATCAAGCCGACCACTTTGTCTATA 420
Qy 421 CGTATCACTTTTTCGCTGTAATATAATGATAGGTTGATGTTGTTGTTCTAAATTTGCG 480
Db 421 CGTATCACTTTTTCGCTGTAATATAATGATAGGTTGATGTTGTTGTTCTAAATTTGCG 480
Qy 481 ACAAATTTCTAAGCCATTAATGACGCAATTAATCAATTAATCAATTAATCAATTAATTA 540
Db 481 ACAAATTTCTAAGCCATTAATGACGCAATTAATCAATTAATCAATTAATCAATTAATTA 540
Qy 541 ATGGCTAAGCGCTTTATCTAAACCGTCTGTCACCATCATCTGTCGTCACCACTTTGTAAT 600
Db 541 ATGGCTAAGCGCTTTATCTAAACCGTCTGTCACCATCATCTGTCGTCACCACTTTGTAAT 600
Qy 601 TTCAATGCTAGTTTCCAAATCAAGAAATCTTGTCTAAAGTTTGTCTACTATTAA 660
Db 601 TTCAATGCTAGTTTCCAAATCAAGAAATCTTGTCTAAAGTTTGTCTACTATTAA 660
Qy 661 AATTGCGTCAATTTGTACACCTCATATACGACATTTTCTAATAGGTAATATTATA 720
Db 661 AATTGCGTCAATTTGTACACCTCATATACGACATTTTCTAATAGGTAATATTATA 720
Qy 721 ATTTTGTATTTTCTA 736
Db 721 ATTTTGTATTTTCTA 736
```

```
RESULT 3
ID AAX90601/c
XX AAX90601 standard; DNA; 2201 BP.
AC AAX90601;
XX
XX
DT 01-OCT-1999 (first entry)
XX
DE Histidine kinase encoding polynucleotide sequence.
XX
KW Histidine kinase; Two component signal transduction system; TCSTS;
KW virulence; bacterial pathogenesis; autophosphorylate; immunisation;
KW antibacterial agent; cognate response regulator; screening;
KW Helicobacter pylori; gastric ulcer; ds.
XX
OS Staphylococcus aureus.
XX
FH Key Location/Qualifiers
CDS 744..2099
FT /*tag= a
FT /product= "Histidine kinase protein"
FT /note= "Has sequence homology to KinA from Lactococcus
FT lactis cremoris"
```

```
XX WO9936508-A1.
XX 22-JUL-1999.
XX
XX 12-JAN-1999; 99WO-US000610.
XX 13-JAN-1998; 98US-00006627.
XX (SMIK ) SMITHLINE BEECHAM CORP.
XX
XX Wallis NG, Shilling LK, Mooney JL, Debouck C, Zhong Y;
XX Jaworski DD, Wang M, Throup JP;
XX
XX WPI; 1999-444390/37.
XX P-PSDB; AAY28601.
XX
XX Novel histidine kinase polynucleotides and polypeptides used to screen
XX for antibacterial compounds.
XX
XX Claim 6; Page 5-6; 43pp; English.
XX
XX The present sequence is a histidine kinase encoding novel polynucleotide
XX isolated from Staphylococcus aureus. The sequence was obtained from a
XX library of chromosomal DNA clones of Staphylococcus aureus in E.coli.
XX Histidine kinase is a component of the two component signal transduction
XX system(TCSTS) which is involved in bacterial pathogenesis and virulence.
XX It undergoes autophosphorylation at a histidine residue and phosphate
XX group is transferred to the cognate response regulator. The nucleotide
XX sequence can be used for diagnosis and staging of diseases, and as
XX reagents for screening genetic mutations. The polynucleotide can also be
XX used for therapeutic or prophylactic purposes and in particular for
XX genetic immunisation. Polypeptides derived from this sequence can be used
XX to produce antibodies and to identify agonists and antagonists which are
XX used to prevent, inhibit or treat diseases, particularly Helicobacter
XX pylori infections, such as gastric ulcers, gastrointestinal carcinoma,
XX and gastritis. The histidine kinase products can be used to screen new
XX antibacterial drugs effective against resistant S.aureus strains
XX
XX Sequence 2201 BP; 863 A; 290 C; 372 G; 676 T; 0 U; 0 Other;
```

```
Query Match 100.0%; Score 736; DB 2; Length 2201;
Best Local Similarity 100.0%; Pred. No. 2.4e-157;
Matches 736; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATTTACGTTTTTGTTCATCGTATCAATACCCAGCCAGCACTGTTTCAATCATTTTGTGTC 60
Db 759 ATTTACGTTTTTGTTCATCGTATCAATACCCAGCCAGCACTGTTTCAATCATTTTGTGTC 700
Qy 61 ACGATCGTATGTTTAACTTTGTTTCGTAATATCTTATATAAATCTACGACATTTGT 120
Db 699 ACGATCGTATGTTTAACTTTGTTTCGTAATATCTTATATAAATCTACGACATTTGT 640
Qy 121 TTCTACTTCATATATATACCCCATACATGATTTAAATTTGTTCCCGTTGCATAACATG 180
Db 639 TTCTACTTCATATATATACCCCATACATGATTTAAATTTGTTCCCGTTGCATAACATG 580
Qy 181 GTTTTATTTTTCAGCTAGATATAGTAAATCATACCTCTGTTTGTGTAATCAATTC 240
Db 579 GTTTTATTTTTCAGCTAGATATAGTAAATCATACCTCTGTTTGTGTAATCAATTC 520
Qy 241 TGGCCATTTACCGTCACCTTTAAAGCGTTCTTATCAATTTGAATACCGTTGACATCAT 300
Db 519 TGGCCATTTACCGTCACCTTTAAAGCGTTCTTATCAATTTGAATACCGTTGACATCAT 460
Qy 301 AATATCCTTTTGTGGCTGACGACGTAATTTGACGAAATTTCTTGCTAAAAGTTCTTCAAT 360
Db 459 AATATCCTTTTGTGGCTGACGACGTAATTTGACGAAATTTCTTGCTAAAAGTTCTTCAAT 400
Qy 361 ATCAAAACGGCTTAACATATATAGTCTGACCGTAATCAAGCCGACCACTTTGTCTATA 420
Db 399 ATCAAAACGGCTTAACATATATAGTCTGACCGTAATCAAGCCGACCACTTTGTCTATA 340
```

QY 421 CGTATCACTTTTCGCTGTAATTAATGATAGGTGTAGATTGTTGTTGTTCTAATTTTGGC 480
 DB |||||
 QY 339 CGTATCACTTTTCGCTGTAATTAATGATAGGTGTAGATTGTTGTTGTTCTAATTTTGGC 280
 DB |||||
 QY 481 ACAAAATTTCTAAGCCATTAATGACGCAACATTAATCTAATATGATTAATTAATCATAGTA 540
 DB |||||
 QY 279 ACAAAATTTCTAAGCCATTAATGACGCAACATTAATCTAATATGATTAATTAATCATAGTA 220
 DB |||||
 QY 541 ATGGCTAAGCGCTTTATCTAAACCGTCTGTGTCATCATCTCTGTGTCACATTTGTAAT 600
 DB |||||
 QY 219 ATGGCTAAGCGCTTTATCTAAACCGTCTGTGTCATCATCTCTGTGTCACATTTGTAAT 160
 DB |||||
 QY 601 TTCAATGTCAGTGTCCCAATCAAGAAATCTTGCTAAGTTTGTTCATCTCTACTATTAA 660
 DB |||||
 QY 159 TTCAATGTCAGTGTCCCAATCAAGAAATCTTGCTAAGTTTGTTCATCTCTACTATTAA 100
 DB |||||
 QY 661 AATTTGGCTCATTTGTACACCTCATATACGACTTTTCTTAATAAGGTAATATATTAA 720
 DB |||||
 QY 99 AATTTGGCTCATTTGTACACCTCATATACGACTTTTCTTAATAAGGTAATATATTAA 40
 DB |||||
 QY 721 AATTTGTATTCTTA 736
 DB |||||
 QY 39 AATTTGTATTCTTA 24
 DB |||||

RESULT 4

AAV75036/c

ID AAV75036 standard; DNA; 1440 BP.

XX AC AAV75036;

XX DT 16-MAR-1999 (first entry)

XX DE Staphylococcus aureus contig SEQ ID #725.

XX KW Computer readable medium; vaccine; S.aureus infection; immunodetection;

XX KW cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;

XX KW skin infection; surgical wound infection; scalded skin syndrome;

XX KW toxic shock syndrome; ds.

XX OS Staphylococcus aureus.

XX FH Key Location/Qualifiers

XX FT misc_feature 1021..1080

XX FT /*tag= a

XX FT /note= "these bases represent a line of missing text in

XX FT the sequence listing in the specification. They are

XX FT included to maintain the nucleotide numbering given in

XX FT the specification for this DNA sequence"

XX PN EP786519-A2.

XX PD 30-JUL-1997.

XX PF 07-JAN-1997; 97EP-00100117.

XX PR 05-JAN-1996; 96US-0009861P.

XX XX (HUMA-) HUMAN GENOME SCI INC.

XX XX Kunsch CA, Choi GH, Barash SC, Dillon PJ, Fannon MR, Rosen CA;

XX XX WPI; 1997-374922/35.

XX XX Polynucleotide(s) and proteins derived from Staphylococcus aureus -

XX XX stored on computer readable medium and used in the production of anti-

XX XX S.aureus vaccines.

XX PS Claim 1; Page 1628-1629; 3271pp; English.

XX XX This sequence represents one of 5191 Staphylococcus aureus DNA sequences

XX XX of the invention. The DNA sequences are recorded on a computer readable

XX XX medium, preferably selected from a floppy or hard disk, random access

CC CC

CC CC

CC CC

CC memory (RAM), read-only memory (ROM) or CD-ROM. Homology searches using
 CC the S.aureus DNA sequences allows putative functions to be assigned so
 CC that protein-encoding or regulatory regions of commercial, therapeutic or
 CC industrial importance can be obtained. Specifically, sequences which are
 CC likely to encode antigens have been identified and these polypeptides can
 CC be used in a vaccine composition against S.aureus infection. The
 CC polypeptides can also be used in a kit for the immunodetection of
 CC S.aureus in a sample. S.aureus is implicated in numerous human diseases,
 CC including cellulitis, eyelid infections, food poisoning, osteomyelitis,
 CC skin and surgical wound infections, scalded skin syndrome, toxic shock
 CC syndrome, etc. Organisms transformed with the DNA sequences can be used
 CC for recombinant production of the polypeptides. The new DNA sequences
 CC (and their fragments) are useful as primers or probes for isolating
 CC homologues of any of the S.aureus DNA sequences contained on the computer
 CC readable medium

XX SQ Sequence 1440 BP; 519 A; 179 C; 241 G; 437 T; 0 U; 64 Other;

Query Match 98.6%; Score 726; DB 2; Length 1440;

Best Local Similarity 99.3%; Pred. No. 4.2e-155;

Matches 726; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 ATTTACGTTTTGTCATCGTATCATACCAACGCCACCACTGTTTCAATCAATTTGTC 60

DB 734 ATTTACGTTTTGTCATCGTATCATACCAACGCCACCACTGTTTCAATCAATTTGTC 675

QY 61 ACGATCGTATGTTTTAACTGTTTCTGTAATATCTTATATAAATCTACGACATTTGT 120

DB 674 ACGATCGTATGTTTTAACTGTTTCTGTAATATCTTATATAAATCTACGACATTTGT 615

QY 121 TTCTACTTCACTATTATAAACCCTATCATGATTTAAATTTGTTCCGTTGCAATACATG 180

DB 614 TTCTACTTCACTATTATAAACCCTATCATGATTTAAATTTGTTCCGTTGCAATACATG 555

QY 181 GTTTTATTTTTCAGCTAGAGATATAGTAATCATACTCTGTTTTTGTAAATTTCAATTC 240

DB 554 GTTTTATTTTTCAGCTAGAGATATAGTAATCATACTCTGTTTTTGTAAATTTCAATTC 495

QY 241 TGGCCATTTACCGTCACTTTAAAGCGTTCTTATCAATTTGTAATACCGTTGACATCGAT 300

DB 494 TGGCCATTTACCGTCACTTTAAAGCGTTCTTATCAATTTGTAATACCGTTGACATCGAT 435

QY 301 AATATCTTTTGTGGCTGACGCTAAATTTGACCAATTTCTTGTAAAGTTTCTTCAAT 360

DB 434 AATATCTTTTGTGGCTGACGCTAAATTTGACCAATTTCTTGTAAAGTTTCTTCAAT 375

QY 361 ATCAAAACGGCTTAACATATATAATCGTCTGCACCGTAATCAAGCCCGCAACTTTGTCATA 420

DB 374 ATCAAAACGGCTTAACATATATAATCGTCTGCACCGTAATCAAGCCCGCAACTTTGTCATA 315

QY 421 CGTATCACTTTTCGCTGTAATTAATGATAGGTGTAGATTGTTGTTGTTCTAATTTTGGC 480

DB 314 CGTATCACTTTTCGCTGTAATTAATGATAGGTGTAGATTGTTGTTGTTCTAATTTTGGC 255

QY 481 ACAAAATTTCTAAGCCATTAATGACGCAACATTAATCTAATATGATTAATTAATCATAGTA 540

DB 254 ACAAAATTTCTAAGCCATTAATGACGCAACATTAATCTAATATGATTAATTAATCATAGTA 195

QY 541 ATGGCTAAGCGCTTTATCTAAACCGTCTGTGTCATCATACTCTGTGTCACATTTGTAAT 600

DB 194 ATGGCTAAGCGCTTTATCTAAACCGTCTGTGTCATCATACTCTGTGTCACATTTGTAAT 135

QY 601 TTCAATGTCAGTGTCCCAATCAAGAAATCTTGCTAAGTTTGTTCATCTCTACTATTAA 660

DB 134 TTCAATGTCAGTGTCCCAATCAAGAAATCTTGCTAAGTTTGTTCATCTCTACTATTAA 75

QY 661 AATTTGCGTCATTTGTACACCTCATATTTACGACTTTTCTTAATAAGGTAATATATTAA 720

DB 74 AATTTGCGTCATTTGTACACCTCATATTTACGACTTTTCTTAATAAGGTAATATATTAA 15

QY 721 AATTTGTATT 731

DB 14 ATANTGNATTT 4

CC polynucleotide sequences from the present invention. AAH55091 to AAH55098
CC represent oligonucleotide sequences and primers which are used in the
CC exemplification of the present invention. N.B. The present invention
CC specifically claims all the polynucleotide sequences given in the
CC sequence listing of the present specification, however the sequence
CC listing only goes up to SEQ ID NO:4454 so even though sequences are given
CC in the disclosure for SEQ ID NO:4465 to 4472, no sequences are present
CC for SEQ ID NO:4455 to 4464

XX Sequence 3099 BP; 1007 A; 488 C; 466 G; 1138 T; 0 U; 0 Other;

Query Match 55.2%; Score 406.4; DB 4; Length 3099;

Best Local Similarity 73.6%; Pred. No. 1.4e-82;

Matches 518; Conservative 0; Mismatches 186; Indels 0; Gaps 0;

Qy 2 TTTACGTTTGTCTATCATACATACCAACGCGCAAGTCTTCAATCAATTTGTCA 61
Db TTTACGTTTGTCTATCATACATACCAACGCGCAAGTCTTCAATCAATTTGTCT 1158
Qy 62 CGATCGTATGGTTTAACTTTTCTTAATATCTTATATACATCTACGACATTTGTT 121
Db TTTAAAGGTTTGAAGTTTAACTTAATAGTAAATACGAATGTAACATCAACGACATTCGTT 1218
Qy 122 TCTACTTCACTATTAATACCCCATACATGATTTAAATTTGTTCCCGTTGCATACATGG 181
Db TCTACTTCACTATTAATACCCCATACATGATTTAAATTTGTTCCCGTTGCATACATGG 1278
Qy 182 TTTTATTTTCAGCTAGAGATATAGTAATCATCTCTGTTTGTGTAATTCATTTCT 241
Db TTAGGATTTTCAGCTAGAGATATAGTAATCATCTCTGTTTGTGTAATTCATTTCT 1338
Qy 242 GCGCATTTACCGTCACTTTAAAGCGTTCTTATCAATTTGTAATACCGTTGACATCGATA 301
Db TGGCCATTAACAGTAACCTTTAAAGCGATCTTATCAATGTAATACCGTTGATATCTAAA 1398
Qy 302 ATATCTTTTGTGCTGACGACGTAATTAATTCGCAATTTCTGTAAGTTCTTCAATA 361
Db AATCTTTTATCTGCTGCTGCGCAATACCGCTTATTTCTGGAGCAATTTCTTATA 1458
Qy 362 TCAACGCGTTAAT 421
Db TCAAGGCGTTTATCAATGTAGTATCTGCGCCATAGTCCAAACCCAGCTACTTTATCATAT 1518
Qy 422 GTATCACTTTTCGCTGTAATATATATATATATATATATATATATATATATATATAT 481
Db GTCTCGCTTTTTCGAGTAAT 1578
Qy 482 CAATTTTCTAAGCCATTAATTTGACGCGCAACATTAATTAATTAATTAATTAATTAAT 541
Db CAATTTTCTAGACCATTTATTTTGGAGCATTAAATTAATTAATTAATTAATTAATTAAT 1638
Qy 542 TGGCTAAGCGCTTTATCTAAACCGTCTGTGCCATCATCTCTGTGCCAATTTGTAATTT 601
Db TTAGATAATGTTTATCCAAACCCACCTTTCCATCATCTCTATATCAACAGTATAATTT 1698
Qy 602 TCATGTGTGAGTTTCAATTTCAAGAACTTTGCTAAGTTTGTGCTATCTTCTACTATATA 661
Db TCATGTGTGAGTTTCAATTTCAAGAACTTTGCTAAGTTTGTGCTATCTTCTACTATATA 1758
Qy 662 ATTTGCGTCAATTTGTACACCTCATATTAACGACTTTTCTTAATAA 705
Db ATATTTGTCAATTTTGACCTCAGCTACATCTTAAACATTAATA 1802

RESULT 7

ABN90912/C

ID ABN90912 standard; DNA; 669 BP.

XX AC ABN90912;

XX DT 24-JUL-2002 (first entry)

XX DE Staphylococcus epidermidis ORF nucleic acid sequence SEQ ID NO:375.

XX Staphylococcus epidermidis; open reading frame; ORF; bacterial infection;
KW antibacterial; gene therapy; gene; ds.
XX Staphylococcus epidermidis.
XX US6380370-B1.
XX 30-APR-2002.
XX 13-AUG-1998; 98US-00134001.
XX 14-AUG-1997; 97US-0055779P.
XX 08-NOV-1997; 97US-0064964P.
XX (GENO-) GENOME THERAPEUTICS CORP.
XX Doucette-Stamm LA, Bush D;
XX WPI; 2002-381255/41.
XX P-PSDB; ABP38367.
XX Novel isolated nucleic acid encoding a Staphylococcus epidermis
XX polypeptide, useful for diagnosing and treating bacterial infections.
XX Disclosure; SEQ ID NO 375; 267pp; English.
XX ABN90538 to ABN93374 represent Staphylococcus epidermidis open reading
XX frame (ORF) nucleic acid sequences which encode the amino acid sequences
XX given in ABP35124 to ABP37960. The S. epidermidis sequences have
XX antibacterial activity and can be used in gene therapy. The sequences can
XX also be used in the diagnosis and treatment of bacterial infections,
XX particularly S. epidermidis infections. The sequences can be used to
XX screen for compounds able to interfere with the S. epidermidis life cycle
XX or inhibit S. epidermidis infection. N.B. The sequence data for this
XX patent did not form part of the printed specification, but was obtained
XX in electronic format directly from the USPTO web site
XX Sequence 669 BP; 248 A; 89 C; 130 G; 202 T; 0 U; 0 Other;

Query Match 54.6%; Score 401.8; DB 6; Length 669;

Best Local Similarity 75.0%; Pred. No. 1.2e-81;

Matches 502; Conservative 0; Mismatches 167; Indels 0; Gaps 0;

Qy 13 TCATCGTATCATATACCAACGCGCAAGTCTTCAATTTGTCAGATCGTATGG 72
Db TCAATCGTATCATATACCAACGCGCAAGTCTTCAATTTGTCAGATCGTATGG 610
Qy 73 TTTTAACTTTGTTTCGTAATATCTTATATAACATCTACGACATTTGTTTCTACTTCAC 132
Db TTTGAGTTTATACGTAATATAACGAATATATAACATCAACGACATTTGTTTCTACTTCAC 550
Qy 133 ATTATAACCCCATACATGATTTAAATTTGTTCCCGTTGCATTAACATGTTTATTTTC 192
Db ATTATACCCCATACGATCGAGATTTGTTCCAGTTGCATGACGTTGTTTACGATTTTC 490
Qy 193 AGCTAGAGATATAGTAATATCATCTCTGTTTGTGTAATTTCTCGGCAATTTAC 252
Db AGCTAGAGATATAGTAATATCATCTCTGTTTGTGTAATTTCTCGGCAATTTAC 430
Qy 253 CGTCACCTTTAAAGCGTTCTTATCAATTTGTAATACGTTGACATCATATATCTTTTC 312
Db AGTAACCTTTAAAGCGATCTTTTATCAATGATATAACCAATGATATCTTAAACATCTTTATC 370
Qy 313 TGGCTGACGCGTAAATTTGTCAGCAATTTCTTGTAAAAGTTCTTCAATATCAACGCGTT 372
Db TGGCTGTCGCGCAATACGCTCTTATTTCTTGAAGCAATTTCTTATATCAAAAGGTTT 310
Qy 373 AACTATATATATCTGTCGACCGTAAATCAAGCCAGCAACTTTGTCTATACGATTCAC 432
Db TACAATGTATGTCATCTGCGCCATAGTCCAAATCCAGCTACTTTTATCATATGTCGCTTT 250
Qy 433 CGCTGTAATTTATAATGATAGTGTAGATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 492

249	CGCAGTAATGATGATAAATGGAGTAGTTGTTTTTGGACGAATTTGCTCACAATTTCTAG	190
Db		
493	GCCATTAATGTGACGGACAATTAATAATCATGATTAATCATGATGAGTAAAGCGC	552
Qy		
189	ACCATTTATGTTTGGAAACATTTAAGTCTTAATATATATAAATCATAGGCTTAGATAATGC	130
Db		
553	TTTATCTAAACCGCTCTGTCCATCATACTCTGTGTCCACATTTGTAATTTTCATGCTGAG	612
Qy		
129	TTTATCTCAAAACCCACCTTTCCATCATCTCTATATCAACAGTATAATTTTCATGAGTTAA	70
Db		
613	TTCCAAATTCACGAATCTTGCTTAAGTTTGTGTCATCTTCTACTATTAAATTTGCGTCAT	672
Qy		
69	CTCAGCTCTATAAATCTAGCAAGATTTTGCTCATCTTCTACAAATTAATAATTTGTCAT	10
Db		
673	TTGTACACC	681
Qy		
9	ATTTCACC	1
Db		

XX	AAH54308/c
XX	AAH54308 standard; DNA; 3230 BP.
XX	AC
XX	AAH54308;
XX	03-SEP-2001 (first entry)
XX	
XX	S. epidermidis genomic polynucleotide sequence SEQ ID NO:3672.
DE	
XX	Staphylococcus epidermidis SRI strain; infection; diagnosis; vaccination;
KW	endocarditis; da.
XX	
XX	Staphylococcus epidermidis.
XX	
PN	WO200134809-A2.
XX	
PD	17-MAY-2001.
XX	
PF	09-NOV-2000; 2000WO-US030782.
XX	
PR	09-NOV-1999; 99US-0164258P.
XX	
PA	(GLAX) GLAXO GROUP LTD.
XX	
PI	Kimmerly WJ;
XX	
DR	WPI; 2001-316495/33.
XX	
PT	Nucleic acids encoding polypeptides from Staphylococcus epidermidis,
PT	useful for vaccinating against infections, e.g. endocarditis.
XX	
PS	Claim 8; Page 1281-1282; 2188pp; English.

AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides (II), given in AAG81454 to AAG83120, from *Staphylococcus epidermidis*. (I) and (II) can have antibacterial activity and therefore can be used in vaccination. The nucleic acids (I) may be used to produce the *S. epidermidis* polypeptides (II) via the production of vectors containing them which are used to produce hosts cells which express the polypeptides. The polypeptides (II) (and/or nucleic acids) may then be used to vaccinate subjects and to raise antibodies against the bacteria. The polypeptides may also be used to assay for antibodies against the bacteria and therefore identify compounds that may be used for the treatment of *S. epidermidis* infections, e.g. endocarditis. AAH53971 to AAH5090 represent specifically claimed *S. epidermidis* genomic DNA polynucleotide sequences from the present invention. AAH55091 to AAH55098 represent oligonucleotide sequences and primers which are used in the exemplification of the present invention. N.B. The present invention specifically claims all the polynucleotide sequences given in the sequence listing of the present specification, however the sequence listing only goes up to SEQ ID NO:4454 so even though sequences are given in the disclosure for SEQ ID NO:4465 to 4472, no sequences are present

[illegible]


```
XX PD 17-MAY-2001.
XX PF 09-NOV-2000; 2000WO-US030782.
XX PR 09-NOV-1999; 99US-0164258P.
XX PA (GLAX ) GLAXO GROUP LTD.
XX PI Kimmery WJ;
XX DR WPI; 2001-316495/33.
XX PT Nucleic acids encoding polypeptides from Staphylococcus epidermidis,
XX FT useful for vaccinating against infections, e.g. endocarditis.
XX PS Claim 8; Page 2015-2016; 2188pp; English.
XX CC AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides
XX CC (II), given in AAG81454 to AAG83120, from Staphylococcus epidermidis. (I)
XX CC and (II) can have antibacterial activity and therefore can be used in
XX CC vaccination. The nucleic acids (I) may be used to produce the S.
XX CC epidermidis polypeptides (II) via the production of vectors containing
XX CC them which are used to produce hosts cells which express the
XX CC polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
XX CC used to vaccinate subjects and to raise antibodies against the bacteria.
XX CC The polypeptides may also be used to assay for other inhibitors of their
XX CC activity and therefore identify compounds that may be used for the
XX CC treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to
XX CC AAH5090 represent specifically claimed S. epidermidis genomic DNA
XX CC polynucleotide sequences from the present invention. AAH5091 to AAH5098
XX CC represent oligonucleotide sequences and primers which are used in the
XX CC exemplification of the present invention. N.B. The present invention
XX CC specifically claims all the polynucleotide sequences given in the
XX CC sequence listing of the present specification, however the sequence
XX CC listing only goes up to SEQ ID NO:4454 so even though sequences are given
XX CC in the disclosure for SEQ ID NO:4465 to 4472, no sequences are present
XX CC for SEQ ID NO:4455 to 4464
XX SQ Sequence 3018 BP; 929 A; 508 C; 459 G; 1122 T; 0 U; 0 Other;

Query Match 47.1%; Score 347; DB 4; Length 3018;
Best Local Similarity 73.9%; Pred. No. 4.2e-69;
Matches 440; Conservative 0; Mismatches 155; Indels 0; Gaps 0;

QY 2 TTTACGTTTGTGTCATCGTATCATACACCCAGCCAGCACTGTTCAATCATTTGTCA 61
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2424 TTGCGCTTAATCATCGAATCACATACCCACGCGTACTGTTCTATGSAATTTTCT 2483

QY 62 CGATCGTATGTTTAACTTGTTCGTAAATATCTTATAAATACATCTACGACATTTGTT 121
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2484 TTATTAAGGTTTGAGTTTATACGTAAATAGCAATGTAACATCAAGCAATTCGTT 2543

QY 122 TCATCTTACATATTAACCCCATACATGATTTAAATTTGTTCCCGTTCATACATGG 181
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2544 TCTACTTCACTATATACCCCATACATGATCGAGATTTGTTTCACTGATGACGTGG 2603

QY 182 TTTTATTTTTCAGTACAGATATAGTAATCATCTGTTTGTGTTTAACTTCTTCT 241
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2604 TTACGATTTTCAGTAAACATATAAATATGCTATTTCTGTTTAGTTAATTTCAATGA 2663

QY 242 GCGCCAATTTACCGTCACTTTAAAGCGTTCTTATCAATTTGTAATCCGTTGACATGATA 301
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2664 TGCCATTACAGTAATCTTAAAGGCACTTTATCAATGATGATATACATGATATCAAA 2723

QY 302 ATATCCTTTTGGCTGACGACGTAAATTTGCAAGAAATTTCTGCTAAAAGTTCTTCAATA 361
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2724 ACATCTTTATCTGGCTGTCTGCGCAATACCGCTCTTATTTCTGCGAGCAATTTCTCTATA 2783

QY 362 TCAACCGCTTAACTATATATCTGTCGACCGCTTAATCAAGCCAGCACTTTGTCTATAC 421
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2784 TCAAGGTTTAAATGATAGTATCATCTGCCCAATAGTCCAAACCCAGCTACTTTATCATAT 2843
```

```
QY 422 GTATCACTTTTCGCTGTAATTAATGATAGGTAGATGTTGTTCTTAATTTTCGGA 481
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2844 GTCTCGCTTTTTCGAGTAATGATGATTAATTTGGAGTAGTGTGTTTTCGACGAATTTGCTA 2903

QY 482 CAAATTTCTAAGCCATTAAATTTGACGCAACATTTAAATCTAATATGATTAATCATAGTAA 541
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2904 CAAATTTCTAGACCAATTTATGTTTGAAGCATTAAAGTCTAAATATATAATCATAGGC 2963

QY 542 TGGCTAAGCGCTTTATCTAAACCGTCTTGTCCATCATCTCTGTGCCACATTTGT 596
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2964 TTAGATAATGCTTTATTCACAAACCCACCTTTCCATCATCTCTATATCAACAGTAT 3018

RESULT 10
AAH53697/c
ID AAH53697 standard; DNA; 501 BP.
XX AC AAH53697;
XX DT 03-SEP-2001 (first entry)
XX DE S. epidermidis open reading frame nucleotide sequence SEQ ID NO:2787.
XX KW Staphylococcus epidermidis SR1 strain; infection; diagnosis; vaccination;
XX OS endocarditis; ds.
XX OS Staphylococcus epidermidis.
XX FN WO200134809-A2.
XX PD 17-MAY-2001.
XX PF 09-NOV-2000; 2000WO-US030782.
XX PR 09-NOV-1999; 99US-0164258P.
XX PA (GLAX ) GLAXO GROUP LTD.
XX PI Kimmery WJ;
XX DR WPI; 2001-316495/33.
XX DR P-PSDB; AAG82847.
XX PT Nucleic acids encoding polypeptides from Staphylococcus epidermidis,
XX PS useful for vaccinating against infections, e.g. endocarditis.
XX CC Claim 8; Page 725; 2188pp; English.
XX CC AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides
XX CC (II), given in AAG81454 to AAG83120, from Staphylococcus epidermidis. (I)
XX CC and (II) can have antibacterial activity and therefore can be used in
XX CC vaccination. The nucleic acids (I) may be used to produce the S.
XX CC epidermidis polypeptides (II) via the production of vectors containing
XX CC them which are used to produce hosts cells which express the
XX CC polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
XX CC used to vaccinate subjects and to raise antibodies against the bacteria.
XX CC The polypeptides may also be used to assay for other inhibitors of their
XX CC activity and therefore identify compounds that may be used for the
XX CC treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to
XX CC AAH5090 represent specifically claimed S. epidermidis genomic DNA
XX CC polynucleotide sequences from the present invention. AAH5091 to AAH5098
XX CC represent oligonucleotide sequences and primers which are used in the
XX CC exemplification of the present invention. N.B. The present invention
XX CC specifically claims all the polynucleotide sequences given in the
XX CC sequence listing of the present specification, however the sequence
XX CC listing only goes up to SEQ ID NO:4454 so even though sequences are given
XX CC in the disclosure for SEQ ID NO:4465 to 4472, no sequences are present
XX CC for SEQ ID NO:4455 to 4464
XX SQ Sequence 501 BP; 185 A; 75 C; 97 G; 144 T; 0 U; 0 Other;

Query Match 41.1%; Score 302.6; DB 4; Length 501;
Best Local Similarity 75.2%; Pred. No. 3.5e-59;
```


Matches 377; Conservative 0; Mismatches 124; Indels 0; Gaps 0;

Qy 13 TCATCGTATCATACCAAGCCAGCAAGTGTTCATATCATTTTGTTCACGATCGTATGG 72
|||
Db 501 TCATCGAATCATACATCCCTACGCCAGTACTGTTCTATGATTTTCTTATTAAAGG 442
|||
Qy 73 TTTTAACCTGTTTGTAAATATCTTATATAAATCAATCTAGCATTTGTTCTACTTCACT 132
|||
Db 441 TTTGAGTTTATTCGTAAATAACGAATGTAACATCAACGACATTCGTTTCTACTTCACT 382
|||
Qy 133 ATTATAACCCCATACATGATTTAAATTTGTTCCGTTTGCATACATCGTTTATTATTTTC 192
|||
Db 381 ATTATACCCCATACATGATTCGAGAAATTTGTTTCAGCTGCATGAGTGTGTTACGATTTTC 322
|||
Qy 193 AGCTAGAAGATATAGTAAATCATACTCTGTTTGTGTTTAAATCAATTTCTGCGCCATTTAC 252
|||
Db 321 AGCTAAACATATAAATCGTATCTGTTTGTGTTTAAATCAATTTCTGCGCCATTTAC 262
|||
Qy 253 CGTCACTTTAAAGGTTCTTATCATTTGTAATACCGTTGACATCGATAATATCTTTTGG 312
|||
Db 261 AGTAACCTTTAAAGGATCTTTTATCAATGATAATACCATTTGATATCTTAAACATCTTTATC 202
|||
Qy 313 TGGCTGACGAGTAAATTTGCAAGAAATCTTGTCTAAAGTTCTTCAATATCAAAAGCGTTT 372
|||
Db 201 TGGCTGTCGCGCAATACCGTCTTATTTCTTGGAGCAATTTCTTATATCAAAAGGTTT 142
|||
Qy 373 AACTATATAATCGTTCACCGTAAATCAAGCCAGCAACTTTTGTTCATACGATCACTTTT 432
|||
Db 141 TACAATGTAGTCACTGCCCCATAGTCCACCCAGTACTTTTATCATATGTCGCTTTT 82
|||
Qy 433 CGCTGTAATTAATGATAGTGTAGATTTGTTGTTCTAAATTTTGGACAAATTTCTTAA 492
|||
Db 81 TGCAGTAAATGATGATAATTTGAGTAGTTGTTTGTGACGAATTTGTTCTACAAATTTCTAG 22
|||
Qy 493 GCCATTAAATGACGCAACAT 513
|||
Db 21 ACCATTATGTTTGGAGCAT 1

RESULT 11

AAH53015/C
ID AAH53015 standard; DNA; 501 BP.
XX
AC AAH53015;
XX
DT 03-SEP-2001 (first entry)
XX
DE S. epidermidis open reading frame nucleotide sequence SEQ ID NO:1423.
XX
KW Staphylococcus epidermidis SRI strain; infection; diagnosis; vaccination;
KW endocarditis; ds.
XX
OS Staphylococcus epidermidis.
XX
FN WO200134809-A2.
XX
PD 17-MAY-2001.
XX
PF 09-NOV-2000; 2000WO-US030782.
XX
PR 09-NOV-1999; 99US-0164258P.
XX
PA (GLAXO) GLAXO GROUP LTD.
XX
PI Kimmerly WJ;
XX
DR WPI; 2001-316495/33.
XX
PT P-PSDB; AAG82165.
XX
PT Nucleic acids encoding polypeptides from Staphylococcus epidermidis,
XX useful for vaccinating against infections, e.g. endocarditis.
PS Claim 8; Page 402; 2189pp; English.

XX
CC AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides
CC (II), given in AAG81454 to AAG83120, from Staphylococcus epidermidis. (I)
CC and (II) can have antibacterial activity and therefore can be used in
CC vaccination. The nucleic acids (I) may be used to produce the S.
CC epidermidis polypeptides (II) via the production of vectors containing
CC them which are used to produce hosts cells which express the
CC polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
CC used to vaccinate subjects and to raise antibodies against the bacteria.
CC The polypeptides may also be used to assay for other inhibitors of their
CC activity and therefore identify compounds that may be used for the
CC treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to
CC AAH55090 represent specifically claimed S. epidermidis genomic DNA
CC polynucleotide sequences from the present invention. AAH55091 to AAH55098
CC represent oligonucleotide sequences and primers which are used in the
CC specification of the present invention. N.B. The present invention
CC specifically claims all the polynucleotide sequences given in the
CC sequence listing of the present specification, however the sequence
CC listing only goes up to SEQ ID NO:4454 so even though sequences are given
CC in the disclosure for SEQ ID NO:4465 to 4472, no sequences are present
CC for SEQ ID NO:4455 to 4464

SQ Sequence 501 BP; 185 A; 75 C; 97 G; 144 T; 0 U; 0 Other;

Query Match 41.1%; Score 302.6; DB 4; Length 501;
Best Local Similarity 75.2%; Pred. No. 3.5e-59;
Matches 377; Conservative 0; Mismatches 124; Indels 0; Gaps 0;

Qy 13 TCATCGTATCATACCAAGCCAGCAAGTGTTCATATCATTTTGTTCACGATCGTATGG 72
|||
Db 501 TCATCGAATCATACATCCCTACGCCAGTACTGTTTCTATGATTTTCTTATTAAAGG 442
|||
Qy 73 TTTTAACCTGTTTGTAAATATCTTATATAAATCTTGTTCACATCTTGTTCACGATCGTATGG 132
|||
Db 441 TTTGAGTTTATTCGTAAATAACGAATGTAACATCAACGACATTCGTTTCTACTTCACT 382
|||
Qy 133 ATTATAACCCCATACATGATTTAAATTTGTTCCGTTTGCATACATCGTTTATTATTTTC 192
|||
Db 381 ATTATACCCCATACATGATTCGAGAAATTTGTTTCAGCTGCATGAGTGTGTTACGATTTTC 322
|||
Qy 193 AGCTAGAAGATATAGTAAATCATACTCTGTTTGTGTTTAAATCAATTTCTGCGCCATTTAC 252
|||
Db 321 AGCTAAACATATAAATCGTATCTGTTTGTGTTTAAATCAATTTCTGCGCCATTTAC 262
|||
Qy 253 CGTCACTTTAAAGGCTTCTTATCAATTTGTAATACCGTTGACATCGATAATATCTTTTG 312
|||
Db 261 AGTAACCTTTAAAGCATCTTTTATCAATGATAATACCATTTGATATCTTAAACATCTTATC 202
|||
Qy 313 TGGCTGACGAGTAAATTTGCACGAATTTCTTGTCTAAAGTTCTTCAATATCAACAGCGTT 372
|||
Db 201 TGGCTGTCGCGCAATACCGCTCTTATTTCTTGGAGCAATTTCTTATATCAAAAGGTTT 142
|||
Qy 373 AACTATATAATCGTTCGACCGTAAATCAAGCCAGCAACTTTTGTTCATACGATCACTTTT 432
|||
Db 141 TACAATGTAGTCACTGCCCCATAGTCCAAACCCAGTACTTTTATCATATGTCGCTTTT 82
|||
Qy 433 CGCTGTAATTAATGATAGTGTAGATTTGTTGTTCTAAATTTTGGACAAATTTCTTAA 492
|||
Db 81 TGCAGTAAATGATGATAATTTGAGTAGTTGTTTGTGACGAATTTGTTCTACAAATTTCTAG 22
|||
Qy 493 GCCATTAAATGACGCAACAT 513
|||
Db 21 ACCATTATGTTTGGAGCAT 1

RESULT 12

ABQ69245_14/C
Continuation (15 of 31) of ABQ69245 from base 1400001 (Listeria innocua DNA sequence #68;
WP Sequence split into 31 fragments LOCUS ABQ69245 Accession ABQ69245
WP Fragment Name Begin End
WP ABQ69245_00 1 110000
WP ABQ69245_01 100001 210000
WP ABQ69245_02 200001 310000

```
WP ABQ69245_03 300001 410000
WP ABQ69245_04 400001 510000
WP ABQ69245_05 500001 610000
WP ABQ69245_06 600001 710000
WP ABQ69245_07 700001 810000
WP ABQ69245_08 800001 910000
WP ABQ69245_09 900001 1010000
WP ABQ69245_10 1000001 1110000
WP ABQ69245_11 1100001 1210000
WP ABQ69245_12 1200001 1310000
WP ABQ69245_13 1300001 1410000
WP ABQ69245_14 1400001 1510000
WP ABQ69245_15 1500001 1610000
WP ABQ69245_16 1600001 1710000
WP ABQ69245_17 1700001 1810000
WP ABQ69245_18 1800001 1910000
WP ABQ69245_19 1900001 2010000
WP ABQ69245_20 2000001 2110000
WP ABQ69245_21 2100001 2210000
WP ABQ69245_22 2200001 2310000
WP ABQ69245_23 2300001 2410000
WP ABQ69245_24 2400001 2510000
WP ABQ69245_25 2500001 2610000
WP ABQ69245_26 2600001 2710000
WP ABQ69245_27 2700001 2810000
WP ABQ69245_28 2800001 2910000
WP ABQ69245_29 2900001 3010000
WP ABQ69245_30 3000001 3011208

Query Match 30.5%; Score 224.8; DB 6; Length 110000;
Best Local Similarity 59.4%; Pred. No. 4.le-41;
Matches 410; Conservative 0; Mismatches 262; Indels 18; Gaps 1;

4 TAGCTTTGTCATCGTATCACATACCCCAAGCCAGCACTGTTTCAATCATTTTGTGCACG 63
10111 TAGTTGTCATGTACGCATCACATACCCAGTCCGCGAACTGTTTGGATGTAATTTCTTC 10052

64 ATCGTATGTTTTAACTTGTTCGTAATATCTTATATAAATCTTATATAAATCTTACGACATTTCTGTC 123
10051 GTCAGGATGATCAATTTTATTTTCGCAAGTACGAACATATACATCTACTTACATTTCTTTC 9992

124 TACTTCACTATTATTAACCCCATACATGATTTAAAATTTGTTCCCGTTGCATACATGTTT 183
9991 TACTTCTGTTTCAATACCCCATACTTTGTTAAGTAACTTCTCCGAGTAAACATAATTT 9932

184 TTTATTTTCAGCTAGAGATATAGTAAATCATACTCTGTTTGTAAATTTGTTAAATCTCAATTTCTGC 243
9931 GACATTTTCCATTAAGTAACTAGTATGTTCTATCTCGCGTTTGTGTTAAATCAATAATTTTC 9872

244 GCCATTTACCGTCACTTTAAAGCGTCTTATCAATTTGTAATACCGTTGACATCGATAAT 303
9871 TTCATCGCGTTTCAATCCGATTTTCTTTTCAACGATTAAAGTTGCGGTATGTAGCGT 9812

304 -----ATCCTTTTGTGGCTGACGACGTAAATTTGCAAGAAATTTCTTGC 345
9811 TGTGTTGTAGCAGATTGTTCTGCAATTTCTACTCGCGTAAATGTAAGCGAGCGTGC 9752

346 TAAAGTTCTTCAATTAACCGCTTAATATATATGCTGTCGCGTAAATCAATTAATCAAGCCC 405
9751 AAGTAGTTCTTCAATTAAGCAATGCGTTAAGCATATAATCATCTGCTCCGTGCGTCAAGTCC 9692

406 AGCAACTTTGTCATAGTATCACTTTTTCGCTGTAATTTAATCATAGGTGTAGATTGTTG 465
9691 AGATAGCGATCAATCACAGATCTCGCGAGTTTCAATGATGGGTGTTGTTTCAC 9632

466 TTGCTTAATTTTCGCAAAATTTCTAAGCCATTAATTTGACGGCAACATTAATAATCTAATAT 525
9631 TTGGCGTACACGACGCAAACTCTACCCGTTTAAATGTTGTAACATTAAGTCTAGTAA 9572

526 GATTAATCATAGTAAATGGCTAAGCGCTTATCTAAACCGTCTGTGCTCCATCATACTCTGT 585
9571 AATAGCATCCCATTTCTTCAATTTAGTCGAGTTCTAGCCAGACGCTCCATCATTAGCAAC 9512
```

```
QY 586 GTCCACATTTGTAATTTTTCATGTGTGAGTTCCTCAATTCAGAAATCTTGTGTAAGTTTGTTC 645
DB 9511 AGCAGTTTCATAATTTTTCGTTGTTGTAATTTCTAGTTCATTAACCGTGTGTAAGTTTTC 9452

QY 646 ATCTTCTACTATTAAAAATTTGCGTCAATTTG 675
DB 9451 ATCTTCTAGGATTAATATTCTATTCTTTG 9422

RESULT 13
ABQ67195_3
Continuation (4 of 5) of ABQ67195 from base 300001 (Listeria innocua contig DNA sequence
WP Sequence split into 5 fragments LOCUS ABQ67195 Accession ABQ67195
WP Fragment Name Begin End
WP ABQ67195_0 110000
WP ABQ67195_1 100001 210000
WP ABQ67195_2 200001 310000
WP ABQ67195_3 300001 410000
WP ABQ67195_4 400001 495269

Query Match 30.5%; Score 224.8; DB 6; Length 110000;
Best Local Similarity 59.4%; Pred. No. 4.le-41;
Matches 410; Conservative 0; Mismatches 262; Indels 18; Gaps 1;

4 TAGCTTTGTCATCGTATCACATACCCCAAGCCAGCACTGTTTCAATCATTTTGTGCACG 63
48840 TAGTTGTCATGTACGCATCACATACCCAGTCCGCGAACTGTTTGGATGTAATTTCTTTC 48899

64 ATCGTATGTTTTAACTTGTTCGTAATATCTTATATAAATCTTATATAAATCTTACGACATTTCTGTC 123
48900 GTCAGGATGATCAATTTTATTTTCGCAAGTACGAACATATACATCTACTTACATTTCTTTC 48959

124 TACTTCACTATTATTAACCCCATACATGATTTAAAATTTGTTCCCGTTGCATACATGTTT 183
48960 TACTTCTGTTTCAATACCCCATACTTTGTTAAGTAACTTCTCCGAGTAAACATAATTT 49019

184 TTTATTTTCAGCTAGAGATATAGTAAATCATACTCTGTTTGTAAATTTGTTAAATCTCAATTTCTGC 243
49020 GACATTTTCCATTAAGTAACTAGTATGTTCTACTCGCGTTTGTGTTAAATCAATAATTTTC 49079

244 GCCATTTACCGTCACTTTAAAGCGTCTTATCAATTTGTAATACCGTTGACATCGATAAT 303
49080 TTCATCGCGTTTCAATCCGATTTTCTTTTCAACGATTAAAGTTGCGGTATGTAGCGT 49139

304 -----ATCCTTTTGTGGCTGACGACGTAAATTTGCAAGAAATTTCTTGC 345
49140 TGTGTTGTAGCAGATTGTTCTGCAATTTTCTACTCGCGTAAATGTAAGCGAGCGTGC 49199

346 TAAAGTTCTTCAATTAACCGCTTAACTATATAATCGTCTGCAACCGTAAATCAAGCCC 405
49200 AAGTAGTTCTTCAATTAAGCAATGCGTTAAGCATATAATCATCTGCTCCGTGCGTCAAGTCC 49259

406 AGCAACTTTGTCATAGTATCACTTTTTCGCTGTAATTTAATCATAGGTGTAGATTGTTG 465
49260 AGATAGCGATCAATCACAGATCTCGCGAGTTTCAATGATGGGTGTTGTTTCAC 49319

466 TTGCTTAATTTTCGCAAAATTTCTAAGCCATTAATTTGACGGCAACATTAATAATCTAATAT 525
49320 TTGGCGTACACGACGCAAACTCTACCCGTTTAAATGTTGTAACATTAAGTCTAGTAA 49379

526 GATTAATCATAGTAAATGGCTAAGCGCTTATCTAAACCGTCTGTGCTCCATCATACTCTGT 585
49380 AATAGCATCCCATTTCTTCAATTTAGTCGAGTTCTAGCCAGACGCTCCATCATTAGCAAC 49439

586 GTCCACATTTGTAATTTTTCATGTGTGAGTTCCAAATTCAGAAATCTTGTGTAAGTTTGTTC 645
49440 AGCAGTTTCATAATTTTTCGTTGTTGTAATTTCTAGTTCATTAACCGTGTGTAAGTTTTC 49499

646 ATCTTCTACTATTAAAAATTTGCGTCAATTTG 675
49500 ATCTTCTAGGATTAATATTCTATTCTTTG 49529
```

RESULT 14
ABA03041_13/c
Continuation (14 of 30) of ABA03041 from base 1300001 (Listeria monocytogenes EGD-e gene
WP Sequence split into 30 fragments LOCUS ABA03041 Accession ABA03041
WP Fragment Name Begin End
WP ABA03041_00 1 110000
WP ABA03041_01 100001 210000
WP ABA03041_02 200001 310000
WP ABA03041_03 300001 410000
WP ABA03041_04 400001 510000
WP ABA03041_05 500001 610000
WP ABA03041_06 600001 710000
WP ABA03041_07 700001 810000
WP ABA03041_08 800001 910000
WP ABA03041_09 900001 1010000
WP ABA03041_10 1000001 1110000
WP ABA03041_11 1100001 1210000
WP ABA03041_12 1200001 1310000
WP ABA03041_13 1300001 1410000
WP ABA03041_14 1400001 1510000
WP ABA03041_15 1500001 1610000
WP ABA03041_16 1600001 1710000
WP ABA03041_17 1700001 1810000
WP ABA03041_18 1800001 1910000
WP ABA03041_19 1900001 2010000
WP ABA03041_20 2000001 2110000
WP ABA03041_21 2100001 2210000
WP ABA03041_22 2200001 2310000
WP ABA03041_23 2300001 2410000
WP ABA03041_24 2400001 2510000
WP ABA03041_25 2500001 2610000
WP ABA03041_26 2600001 2710000
WP ABA03041_27 2700001 2810000
WP ABA03041_28 2800001 2910000
WP ABA03041_29 2900001 2944528

Query Match 27.5%; Score 202.2; DB 6; Length 110000;
Best Local Similarity 57.8%; Pred. No. 5.5e-36;
Matches 391; Conservative 0; Mismatches 268; Indels 18; Gaps 1;
Oy 17 CGTATCATACCCAGCGACTGTTTCAATCATTTTGTACGATCGTATGGTTT 76
Db 103223 CGCATCATACCTGTCGCGAACTGTTTGAATATAAATCTTCTCTCGTAGGATGATCA 103164
Oy 77 AACTTGTTCGTAATATCTTATATAAATCTACGACATTTGTTTCTACTTCACTATTA 136
Db 103163 ATTTTATTCGTAAGTACGACGATACATCCACTACATTTGTTCTACTTCTGTTCA 103104
Oy 137 TAACCCCATACATGATTTAAATTTGTTCCCGTTCGATCAACATGTTTATTTTCAGCT 196
Db 103103 TAACCCCATACATTTTATTTGATGTAACACTTCCCGCGTAAGACGATATTAACATTTTCCATT 103044
Oy 197 AGAAGATATAGTAATATCTGTTTGTGTTTGTAAATTCATTTCTGC----- 243
Db 103043 AAGTAAGTAAAGTTTCACTCTCGTTTGTGTTTGTAAATTCATTTCTGTCGCGTTG 102984
Oy 244 -----GCCATTTACCGTCACCTTTAAAGCGTTCTTATCAATTTGTAATACCGTTGACATCG 298
Db 102983 ACAATGCGATTTTCTTTTCAACATTTAGATTACGATCTAGCGTTGTTGTTAGCA 102924
Oy 299 ATAATATCTTTTGTGCGTACGAGTAAATTTGACGAATTTCTTGTCTTAAAGTTTCTTCA 358
Db 102923 GATTGTTCTGCAATTTTCCACCGCGCAATAGCGAGCGAAGCGCGAAGTAATTTCTTCA 102864
Oy 359 ATATCAACGCGTTAACTATATATCTCTCAGCGTAAATCAAGCCAGCAACTTTGTGCA 418
Db 102863 ATAGCGAAAGGTTTGACGATGTAATCATCTGCTCCGTGATCCAGTCCGGATACACGATCG 102804
Oy 419 TACGTATCACTTTTTCGCTGTAATATATATAGTGTAGATGTTGTTGTTCTAATTTTG 478
Db 102803 ATACAGAGTCAGTCCGCGTTTATCAATAATAGGTTGTTGTTTCACTTTCGCGCAGCA 102744
Oy 479 CGACAAATTTCTAAGCCATTAATTTGACGCGCAACATTTAAATCTAATATGATTAATCATAG 538

Db 102743 CGACAAACTTCTACCCCGTTTAAATGTGGCAACATTAGATCGAGTAAATATAGCATCCCAT 102684
Oy 539 TAATGGCTAAGCGGTTTATCTTAAACCGCTTGTCCATCATACTCTGTGTCCACATGTAA 598
Db 102683 TCTTCATTTAAGTGGAGTTTCCAGTCCAGCGGTCATCATTTAGCAACCGCTGTTTCATAA 102624
Oy 599 TTTTCATGTGTGAGTTCCTCAATTTCAAGAAATCTTGTAAAGTTTGTTCATCTTCTACTATT 658
Db 102623 TTTTCATGCTGGAGTTCTAGTTTCAATAAAGCGTGTCTAAGTTTTCATCTTCTACGATT 102564
Oy 659 AAAATTTGCGTCATTTG 675
Db 102563 AGTATTTCTATTTCATTTG 102547

RESULT 15
ABA03041_14/c
Continuation (15 of 30) of ABA03041 from base 1400001 (Listeria monocytogenes EGD-e gene
WP Sequence split into 30 fragments LOCUS ABA03041 Accession ABA03041
WP Fragment Name Begin End
WP ABA03041_00 1 110000
WP ABA03041_01 100001 210000
WP ABA03041_02 200001 310000
WP ABA03041_03 300001 410000
WP ABA03041_04 400001 510000
WP ABA03041_05 500001 610000
WP ABA03041_06 600001 710000
WP ABA03041_07 700001 810000
WP ABA03041_08 800001 910000
WP ABA03041_09 900001 1010000
WP ABA03041_10 1000001 1110000
WP ABA03041_11 1100001 1210000
WP ABA03041_12 1200001 1310000
WP ABA03041_13 1300001 1410000
WP ABA03041_14 1400001 1510000
WP ABA03041_15 1500001 1610000
WP ABA03041_16 1600001 1710000
WP ABA03041_17 1700001 1810000
WP ABA03041_18 1800001 1910000
WP ABA03041_19 1900001 2010000
WP ABA03041_20 2000001 2110000
WP ABA03041_21 2100001 2210000
WP ABA03041_22 2200001 2310000
WP ABA03041_23 2300001 2410000
WP ABA03041_24 2400001 2510000
WP ABA03041_25 2500001 2610000
WP ABA03041_26 2600001 2710000
WP ABA03041_27 2700001 2810000
WP ABA03041_28 2800001 2910000
WP ABA03041_29 2900001 2944528

Query Match 27.5%; Score 202.2; DB 6; Length 110000;
Best Local Similarity 57.8%; Pred. No. 5.5e-36;
Matches 391; Conservative 0; Mismatches 268; Indels 18; Gaps 1;
Oy 17 CGTATCATACCCAGCGACTGTTTCAATCATTTTGTACGATCGTATGGTTT 76
Db 3223 CGCATCATACCTGTCGCGAACTGTTTGAATATAAATCTTCTCTCGTAGGATGATCA 3164
Oy 77 AACTTGTTCGTAATATCTTATATAAATCTACGACATTTGTTTCTACTTCACTATTA 136
Db 3163 ATTTTATTCGTAAGTAAACGATACATCCACTACATTTGTTTCTACTTCTGTTTCA 3104
Oy 137 TAACCCCATACATGATTTAAATTTGTTCCCGTTCGATCAACATGTTTATTTTCAGCT 196
Db 3103 TAACCCCATACATTTTATTTGAGTAAACACTTCCCGCGTAAGACGATATTAACATTTTCCATT 3044
Oy 197 AGAAGATATAGTAATATCTGTTTGTGTTTGTAAATTCATTTCTGC----- 243
Db 3043 AAGTAAGTAAAGTTTCACTCTCGTTTGTGTTTGTAAATTCATTTCTGTCGCGTTG 2984
Oy 244 -----GCCATTTACCGTCACCTTTAAAGCGTTCTTATCAATTTGTAATACCGTTGACATCG 298

```
Db 2983 ACAATGCGATTTTCCTTTTCAACAAATTAGATTACGATACCTGTAGCGTTGTTGTTTAGCA 2924
Qy 299 ATAATATCCTTTTGTGGCTGACGAGTAATAATTGCACGAATTTCTTGCTAAAAGTTCTTCA 358
Db 2923 GATTGTTCTGCAATTTTCCACCCGACGCAATAGCGAGCGAAGGGCGCGCAAGTAATCTTCA 2864
Qy 359 ATATCAAAAGGCTTAACATATATAATCGTCTGCACCGTAATCAAGCCGACGCAACTTTGTCA 418
Db 2863 ATAGCGAAAGTTTGACGATGTAATCATCTGCTCCGGATCCAGTCCGGATACAGATCG 2804
Qy 419 TAGGTATCACTTTTTCGCTGTAATTAATAATGATAGGTGTAGATTGTTGTTGTTCTAATTTTG 478
Db 2803 ATACAGAGTCACGTGGGTTATCATAATAATGGGTTTGTGTTTCACTTGGCGCACACGA 2744
Qy 479 CGACAAATTTCTAAGCCATTAATTGACGGCAACATTAATCTAATATGATTAATAATCATAG 538
Db 2743 CGACAAACTTCTACCCCGTTTAAATGTGGCAACATTAGATCGAGTAAATAGCATCCCAT 2684
Qy 539 TAATGGCTAAGCGCTTTATCTAAACGCTCTGTCCCATCACTCTGTGTCCACATTGTAA 598
Db 2683 TCTTCATTAGTGGAGTTGAGTCCAGCGCGTCCATCATAGCAACCGCTGTTTCATAA 2624
Qy 599 TTTTCATGTGTGAGTTCCAAATTCAGAAATCTTGCTAAGTTTGTGTTTCATCTTCTTCTATT 658
Db 2623 TTTTCATGTGTGAGTTCTAGTTCATAAAGGCTAAGTTTTTTTCATCTTCTAGATT 2564
Qy 659 AAAATTGCGTCATTG 675
Db 2563 AGTATTCATTTCATTG 2547
```

Search completed: October 5, 2004, 03:49:28
Job time : 305.968 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 4, 2004, 17:52:23 ; Search time 57.5284 Seconds
(without alignments)
1075.607 Million cell updates/sec

Title: US-09-006-627-4
Perfect score: 1122
Sequence: 1 MTQILLVEDEQNLFARLEFE.....KPYDRDKMIEVTRGVGVIR 219

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*

1: geneseqp1980s:*

2: geneseqp1990s:*

3: geneseqp2000s:*

4: geneseqp2001s:*

5: geneseqp2002s:*

6: geneseqp2003as:*

7: geneseqp2003bs:*

8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1122	100.0	219	2	Aaw19274
2	1122	100.0	219	2	Aay28602
3	1122	100.0	219	6	Abm73031
4	968	86.3	222	5	Abp38367
5	760	67.7	166	4	Ag82165
6	760	67.7	166	4	Ag82165
7	654	58.3	228	7	Adc96909
8	628	56.0	226	5	Abb48293
9	592.5	52.8	229	6	Abd09376
10	583	52.0	230	5	Abb54940
11	574.5	51.2	229	5	Abp27835
12	574	51.2	228	5	Abp27042
13	498	44.4	230	6	Abu25596
14	490.5	43.7	232	6	Abu23640
15	482.5	43.0	238	6	Abd08750
16	476	42.4	232	6	Abu24519
17	476	42.4	238	5	Abb48983
18	461.5	41.1	226	4	Ag90586
19	461.5	41.1	226	5	Abb83914
20	458	40.8	246	7	Adc94508
21	458	40.8	261	6	Abu29650
22	457	40.7	232	4	Ag90722
23	456.5	40.7	235	6	Abu18458
24	456	40.6	234	4	Abu35155
25	456	40.6	234	6	Abu14636

26	452	40.3	233	4	AAU33417	Aau33417 Enterococ
27	452	40.3	246	5	ABP38330	Abp38330 Staphyloc
28	449.5	40.1	234	7	ADC97119	Adc97119 E. faeciu
29	449	40.0	233	2	AAV06430	Aay06430 Staphyloc
30	449	40.0	233	3	AAB15430	Abi15430 S. aureus
31	449	40.0	233	4	AAU34180	Aau34180 Staphyloc
32	449	40.0	233	4	AAU00850	Aau00850 S. aureus
33	449	40.0	233	4	AAAG81693	Ag81693 S. epider
34	449	40.0	233	6	ABU08234	Abu08234 Staphyloc
35	449	40.0	234	5	ABP38218	Abp38218 Staphyloc
36	449	40.0	235	4	AAU37407	Aau37407 Staphyloc
37	449	40.0	235	4	AAU37081	Aau37081 Staphyloc
38	449	40.0	235	6	ABU16353	Abu16353 Protein e
39	449	40.0	235	6	ABM72823	Abm72823 Staphyloc
40	449	40.0	239	2	AAW89791	Aaw89791 Staphyloc
41	449	40.0	239	6	ABU42897	Abu42897 Protein e
42	447	39.8	236	7	ADB80241	Adb80241 Mycobacte
43	447	39.8	238	6	ABU43704	Abu43704 Protein e
44	444.5	39.6	256	5	ABP65455	Abp65455 Bifidobac
45	442.5	39.4	228	6	ABU35827	Abu35827 Protein e

ALIGNMENTS

RESULT 1

ID	AAW19274	standard; protein; 219 AA.
XX	AAW19274;	
AC	AAW19274;	
XX		
DT	17-OCT-2003 (revised)	
DT	16-JAN-1998 (first entry)	
XX		
DE	Staphylococcus aureus novel response regulator protein.	
XX		
KW	Antibacterial; anti-Staphylococcus; gene therapy; bacterial infection; prophylaxis; two component signal transduction system; TCSTG; ResD;	
KW	response regulator; Staphylococcus; Bacillus; vaccine; wound treatment; dental; surgery; antibody.	
XX		
OS	Staphylococcus aureus; WCUH29.	
XX		
PN	WO9723506-A1.	
XX		
PD	03-JUL-1997.	
XX		
PF	20-DEC-1996; 96WO-GB003261.	
XX		
PR	22-DEC-1995; 95GB-00026359.	
XX		
PA	(SMIK) SMITHKLINE BEECHAM PLC.	
XX		
PI	Wallis N, Hodgson JE;	
XX		
DR	WPI; 1997-350970/32.	
XX		
DR	N-PSDB; AAT72328.	
XX		
PT	Isolated nucleic acid encoding Staphylococcus aureus response regulator - useful to treat or prevent bacterial infections, in vaccines and for diagnosis.	
XX		
PS	Claim 12; Page 46; 59pp; English.	
XX		
CC	The present sequence represents a novel response regulator protein from Staphylococcus aureus. This polypeptide is a response regulatory component of the bacterial two component signal transduction system (TCSTG), related to the ResD protein of Bacillus subtilis (42% identity).	
CC	Inhibition of the polypeptide prevents the bacteria from establishing or maintaining infection by preventing them from producing the factors necessary for pathogenesis. The DNA encoding this protein can be used to produce recombinant polypeptides in vivo and in gene therapy. The antagonist and agonists can also be used to inhibit bacterial infection.	

CC Typical applications are in wound treatment (to prevent bacterial
 CC adherence), prophylaxis in dental or other surgery (as an alternative or
 CC adjunct to antibiotic therapy) or for treating in-dwelling devices
 CC immediately before insertion. Detection of the present polypeptide or the
 CC encoding polynucleotide by usual hybridisation or immunoassays, can be
 CC used to diagnose infection. Elevated levels of expression would indicate
 CC bacterial infection. The polypeptide can also be used in protective
 CC vaccines to induce antibody production. The antibodies themselves can be
 CC used as a diagnostic reagent or as a therapeutic antagonist.. (Updated on
 CC 17-OCT-2003 to standardise OS field)
 XX
 SQ Sequence 219 AA;

Query Match 100.0%; Score 1122; DB 2; Length 219;
 Best Local Similarity 100.0%; Pred. No. 6.6e-102;
 Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MTQILIVEDQNLFLELELTHENVNVDTEYDQGLDKALSHYVDLIIDMLPSING 60
 DB 1 MTQILIVEDQNLFLELELTHENVNVDTEYDQGLDKALSHYVDLIIDMLPSING 60
 QY 61 LEICRKIROQOSTPIIIITAKSDTYDKVAGLDYGADYIVKPFIDIELLARIRAILRRQP 120
 DB 61 LEICRKIROQOSTPIIIITAKSDTYDKVAGLDYGADYIVKPFIDIELLARIRAILRRQP 120
 QY 121 OKDIIDVNGITIDKNAPKVTNGAEIELTKTEYDLYLLAENKNHVMQREQILNHVWGYN 180
 DB 121 OKDIIDVNGITIDKNAPKVTNGAEIELTKTEYDLYLLAENKNHVMQREQILNHVWGYN 180
 QY 181 SEVENTNVVDVYIRLNKLPYDRDKMIETVRGVGVIR 219
 DB 181 SEVENTNVVDVYIRLNKLPYDRDKMIETVRGVGVIR 219

RESULT 2
 AAY28602
 ID AAY28602 standard; protein; 219 AA.
 AC AAY28602;
 XX
 XX 01-OCT-1999 (first entry)
 XX
 XX Cognate response regulator of Histidine kinase.
 XX
 XX Cognate response regulator; Two component signal transduction system;
 KW TCSTS; Histidine kinase; virulence; bacterial pathogenesis; screening;
 KW autophosphorylate; antibacterial agent; Helicobacter pylori.
 XX
 XX Staphylococcus aureus.
 XX
 XX WO9936508-A1.
 XX
 XX 22-JUL-1999.
 XX
 XX 12-JAN-1999; 99WO-US000610.
 XX
 XX 13-JAN-1998; 98US-00006627.
 XX
 XX (SMIK) SMITHKLINE BEECHAM CORP.
 XX
 XX Wallis NG, Shilling LK, Mooney JL, Debouck C, Zhong Y;
 PI Jaworecki DD, Wang M, Throup JP;
 XX
 XX WPI; 1999-444390/37.
 DR N-PSDB; AAX90602.
 XX
 XX Novel histidine kinase polynucleotides and polypeptides used to screen
 PT for antibacterial compounds.
 XX
 XX Claim 23; Page 7; 43pp; English.
 PS
 XX The present sequence is the Staphylococcus aureus response regulator
 CC cognate to histidine kinase. Histidine kinase is a component of the two
 CC

CC component signal transduction systems (TCSTS) which is involved in
 CC bacterial pathogenesis and virulence. It undergoes autophosphorylation at
 CC a histidine residue and the phosphate group is transferred to the cognate
 CC response regulator. This novel response regulator shows 42% identity to
 CC the ResD response regulator protein from Bacillus subtilis. The cognate
 CC response regulator can be used for screening drugs which interfere with
 CC the interaction of histidine kinase and the response regulator. The
 CC histidine kinase products can be used to screen and identify new
 CC antibacterial drugs, agonists and antagonists effective against
 CC Helicobacter pylori infections and infections caused by resistant
 CC S.aureus strains
 XX
 SQ Sequence 219 AA;

Query Match 100.0%; Score 1122; DB 2; Length 219;
 Best Local Similarity 100.0%; Pred. No. 6.6e-102;
 Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MTQILIVEDQNLFLELELTHENVNVDTEYDQGLDKALSHYVDLIIDMLPSING 60
 DB 1 MTQILIVEDQNLFLELELTHENVNVDTEYDQGLDKALSHYVDLIIDMLPSING 60
 QY 61 LEICRKIROQOSTPIIIITAKSDTYDKVAGLDYGADYIVKPFIDIELLARIRAILRRQP 120
 DB 61 LEICRKIROQOSTPIIIITAKSDTYDKVAGLDYGADYIVKPFIDIELLARIRAILRRQP 120
 QY 121 OKDIIDVNGITIDKNAPKVTNGAEIELTKTEYDLYLLAENKNHVMQREQILNHVWGYN 180
 DB 121 OKDIIDVNGITIDKNAPKVTNGAEIELTKTEYDLYLLAENKNHVMQREQILNHVWGYN 180
 QY 181 SEVENTNVVDVYIRLNKLPYDRDKMIETVRGVGVIR 219
 DB 181 SEVENTNVVDVYIRLNKLPYDRDKMIETVRGVGVIR 219

RESULT 3
 ABM73031
 ID ABM73031 standard; protein; 219 AA.
 XX
 AC ABM73031;
 XX
 XX 20-NOV-2003 (first entry)
 XX
 XX Staphylococcus aureus protein #2271.
 XX
 XX Antibacterial; vaccine; gene therapy; infection; sepsis; diagnosis;
 KW enzymatic assay; antibiotic target.
 XX
 XX Staphylococcus aureus.
 XX
 XX WO200294868-A2.
 XX
 XX 28-NOV-2002.
 XX
 XX 27-MAR-2002; 2002WO-IB002637.
 XX
 XX 27-MAR-2001; 2001GB-00007661.
 XX
 XX (CHIR-) CHIRON SPA.
 XX
 XX Masignani V, Mora M, Scarselli M;
 XX
 XX WPI; 2003-120786/11.
 DR N-PSDB; ACF74591.
 XX
 XX New Staphylococcus aureus protein, useful as a vaccine for treating or
 PT preventing Staphylococcal infection, specifically an infection caused by
 PT S. aureus, e.g. sepsis.
 XX
 XX Claim 1; SEQ ID NO 4542; 49pp; English.
 PS
 XX The invention relates to novel genes and encoded proteins from
 CC Staphylococcus aureus. A composition comprising the S. aureus protein, a
 CC

CC nucleic acid encoding the protein, or an antibody to the protein, is
 CC preventing infection due to Staphylococcus epidermidis, specifically an
 CC infection caused by *S. aureus*. The composition is particularly useful for
 CC treating or preventing sepsis in a patient. The composition can also be
 CC used for diagnostics. The protein is also used in an assay for enzymatic
 CC studies and as a target for antibiotics. This sequence represents one of
 CC the novel *S. aureus* proteins of the invention
 XX
 SQ Sequence 219 AA;

Query Match 100.0%; Score 1122; DB 6; Length 219;
 Best Local Similarity 100.0%; Pred. No. 6.6e-102; Indels 0; Gaps 0;
 Matches 219; Conservative 0; Mismatches 0;
 Qy 1 MTQLIVDEQNARFLLELTHENVNVDTEYDQGLDKALSHYDYLIIIDMLPSSING 60
 Db 1 MTQLIVDEQNARFLLELTHENVNVDTEYDQGLDKALSHYDYLIIIDMLPSSING 60
 Qy 61 LEICRKIRQQOSTPIIIITAKSDTYKVGADYDADDYIVKPFDEIELLARIRAILRRQP 120
 Db 61 LEICRKIRQQOSTPIIIITAKSDTYKVGADYDADDYIVKPFDEIELLARIRAILRRQP 120
 Qy 121 QKDIIDVNGITIDKNAFKVTNGAEIELTKTEYDLYLLAENKHNVMQREQILNHWGYN 180
 Db 121 QKDIIDVNGITIDKNAFKVTNGAEIELTKTEYDLYLLAENKHNVMQREQILNHWGYN 180
 Qy 181 SEVTNVVDVYIRYLNRKLKPYDRDKMIETVRGVGYVIR 219
 Db 181 SEVTNVVDVYIRYLNRKLKPYDRDKMIETVRGVGYVIR 219

RESULT 4
 ABP38367
 ID ABP38367 standard; protein; 222 AA.
 XX
 AC ABP38367;
 XX
 DT 24-JUL-2002 (first entry)
 XX
 DE Staphylococcus epidermidis ORF amino acid sequence SEQ ID NO:3212.
 XX
 KW Staphylococcus epidermidis; open reading frame; ORF; bacterial infection;
 KW antibacterial; gene therapy.
 XX
 OS Staphylococcus epidermidis.
 XX
 PN US6380370-B1.
 XX
 PD 30-APR-2002.
 XX
 PF 13-AUG-1998; 98US-00134001.
 XX
 PR 14-AUG-1997; 97US-0055779P.
 PR 08-NOV-1997; 97US-0064964P.
 XX
 XX (GENO-) GENOME THERAPEUTICS CORP.
 XX
 PI Doucette-Stamm LA, Bush D;
 XX
 DR WPI; 2002-381255/41.
 DR N-PSDB; ABN90912.
 XX

Novel isolated nucleic acid encoding a Staphylococcus epidermidis
 PT polypeptide, useful for diagnosing and treating bacterial infections.
 XX
 PS Disclosure; SEQ ID NO 3212; 267pp; English.
 XX
 CC ABN90538 to ABN93374 represent Staphylococcus epidermidis open reading
 CC frame (ORF) nucleic acid sequences which encode the amino acid sequences
 CC given in ABP35124 to ABP37960. The *S. epidermidis* sequences have
 CC antibacterial activity and can be used in gene therapy. The sequences can
 CC also be used in the diagnosis and treatment of bacterial infections,

CC particularly *S. epidermidis* infections. The sequences can be used to
 CC screen for compounds able to interfere with the *S. epidermidis* life cycle
 CC or inhibit *S. epidermidis* infection. N.B. The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from the USPTO web site
 XX

SQ Sequence 222 AA;

Query Match 86.3%; Score 968; DB 5; Length 222;
 Best Local Similarity 84.5%; Pred. No. 9.2e-87;
 Matches 185; Conservative 21; Mismatches 13; Indels 0; Gaps 0;
 Qy 1 MTQLIVDEQNARFLLELTHENVNVDTEYDQGLDKALSHYDYLIIIDMLPSSING 60
 Db 4 MTNLIIVDEQNARFLLELTHENVNVDTEYDQGLDKALSKPYDLYIILMLPNING 63
 Qy 61 LEICRKIRQQOSTPIIIITAKSDTYKVGADYDADDYIVKPFDEIELLARIRAILRRQP 120
 Db 64 LEICRKIRQQOSTPIIIITAKSETYDKVAGLDYDADDYIVKPFDEIELLARIRAILRRQP 123
 Qy 121 QKDIIDVNGITIDKNAFKVTNGAEIELTKTEYDLYLLAENKHNVMQREQILNHWGYN 180
 Db 124 DKQVLDINGIIDKDAFKVTNGHQLTEYDLYLLAENKHNVMQREQILNHWGYN 183
 Qy 181 SEVTNVVDVYIRYLNRKLKPYDRDKMIETVRGVGYVIR 219
 Db 184 SEVTNVVDVYIRYLNRKLKPFNKEKSIETVRGVGYVIR 222

RESULT 5
 AAG82165
 ID AAG82165 standard; protein; 166 AA.
 XX
 AC AAG82165;
 XX
 DT 03-SEP-2001 (first entry)
 XX
 DE S. epidermidis open reading frame protein sequence SEQ ID NO:1424.
 XX
 KW Staphylococcus epidermidis SRI strain; infection; diagnosis; vaccination;
 KW endocarditis.
 XX
 OS Staphylococcus epidermidis.
 XX
 PN WO200134809-A2.
 XX
 PD 17-MAY-2001.
 XX
 PF 09-NOV-2000; 2000WO-US030782.
 XX
 PR 09-NOV-1999; 99US-0164258P.
 XX
 PA (GLAX) GLAXO GROUP LTD.
 XX
 PI Kimmerly WJ;
 XX
 DR WPI; 2001-316495/33.
 DR N-PSDB; AAG83015.
 XX
 PT Nucleic acids encoding polypeptides from Staphylococcus epidermidis,
 XX useful for vaccinating against infections, e.g. endocarditis.
 PS Claim 18; Page 402; 2188pp; English.

AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides
 CC (II), given in AAG81454 to AAG83120, from Staphylococcus epidermidis. (I)
 CC and (II) can have antibacterial activity and therefore can be used in
 CC vaccination. The nucleic acids (I) may be used to produce the *S.*
 CC epidermidis polypeptides (II) via the production of vectors containing
 CC them which are used to produce hosts cells which express the
 CC polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
 CC used to vaccinate subjects and to raise antibodies against the bacteria.
 CC The polypeptides may also be used to assay for other inhibitors of their

CC activity and therefore identify compounds that may be used for the
 CC treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to
 CC AAH5090 represent specifically claimed S. epidermidis genomic DNA
 CC polynucleotide sequences from the present invention. AAH55091 to AAH55098
 CC represent oligonucleotide sequences and primers which are used in the
 CC exemplification of the present invention. N.B. The present invention
 CC specifically claims all the polynucleotide sequences given in the
 CC sequence listing of the present specification, however the sequence
 CC listing only goes up to SEQ ID NO:4454 so even though sequences are given
 CC in the disclosure for SEQ ID NO:4465 to 4472, no sequences are present
 CC for SEQ ID NO:4455 to 4464
 XX
 SQ Sequence 166 AA;

Query Match 67.7%; Score 760; DB 4; Length 166;
 Best Local Similarity 85.5%; Pred. No. 1.7e-66;
 Matches 142; Conservative 19; Mismatches 5; Indels 0; Gaps 0;
 QY 54 MLPSINGLEICRKIROQOSTPIIIITAKSDTYDKVAGLDYGADYIVKPFDELLARIR 113
 DB 1 MLFNINGLEICRQIRQKTTPIIITAKSETYDKVAGLDYGADYIVKPFDELLARIR 60
 QY 114 AILRRQPKDIIIDVNGITIDKNAFKVTVNGAEIELTKTEYDLYLLAENKNHVMQREQIL 173
 DB 61 AVLRRQPKDVLIDINGIIIDKDAFKVTVNGHQLELTKEYDLYLLAENKNHVMQREQIL 120
 QY 174 NHWGYNSEVTNVVDVYIYRLNKLKPYDRDKMIETVRGVGVIR 219
 DB 121 DHWGYNSEVTNVVDVYIYRLNKLKPFNKEKSIETVRGVGVIR 166

RESULT 6
 AAG82847
 ID AAG82847 standard; protein; 166 AA.

XX AC AAG82847;
 XX DT 03-SEP-2001 (first entry)
 XX DE S. epidermidis open reading frame protein sequence SEQ ID NO:2788.
 XX KW Staphylococcus epidermidis SR1 strain; infection; diagnosis; vaccination;
 XX KW endocarditis.
 XX OS Staphylococcus epidermidis.
 XX PN WO200134809-A2.
 XX PD 17-MAY-2001.
 XX PF 09-NOV-2000; 2000WO-US030782.
 XX PR 09-NOV-1999; 99US-0164258P.
 XX PA (GLAX) GLAXO GROUP LTD.
 XX PI Kimmberly WJ;
 XX WPI; 2001-316495/33.
 XX N-PSDB; AAH53697.
 XX Nucleic acids encoding polypeptides from Staphylococcus epidermidis,
 XX useful for vaccinating against infections, e.g. endocarditis.
 XX Claim 18; Page 725; 2188pp; English.

XX AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides
 XX (II) given in AAG81454 to AAG8120, from Staphylococcus epidermidis. (I)
 XX and (II) can have antibacterial activity and therefore can be used in
 XX vaccination. The nucleic acids (I) may be used to produce the S.
 XX epidermidis polypeptides (II) via the production of vectors containing
 XX them which are used to produce hosts cells which express the
 XX polypeptides. The polypeptides (II) (and/or nucleic acids) may then be

CC used to vaccinate subjects and to raise antibodies against the bacteria.
 CC The polypeptides may also be used to assay for other inhibitors of their
 CC activity and therefore identify compounds that may be used for the
 CC treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to
 CC AAH5090 represent specifically claimed S. epidermidis genomic DNA
 CC polynucleotide sequences from the present invention. AAH55091 to AAH55098
 CC represent oligonucleotide sequences and primers which are used in the
 CC exemplification of the present invention. N.B. The present invention
 CC specifically claims all the polynucleotide sequences given in the
 CC sequence listing of the present specification, however the sequence
 CC listing only goes up to SEQ ID NO:4454 so even though sequences are given
 CC in the disclosure for SEQ ID NO:4465 to 4472, no sequences are present
 CC for SEQ ID NO:4455 to 4464
 XX
 SQ Sequence 166 AA;

Query Match 67.7%; Score 760; DB 4; Length 166;
 Best Local Similarity 85.5%; Pred. No. 1.7e-66;
 Matches 142; Conservative 19; Mismatches 5; Indels 0; Gaps 0;
 QY 54 MLPSINGLEICRKIROQOSTPIIIITAKSDTYDKVAGLDYGADYIVKPFDELLARIR 113
 DB 1 MLFNINGLEICRQIRQKTTPIIITAKSETYDKVAGLDYGADYIVKPFDELLARIR 60
 QY 114 AILRRQPKDIIIDVNGITIDKNAFKVTVNGAEIELTKTEYDLYLLAENKNHVMQREQIL 173
 DB 61 AVLRRQPKDVLIDINGIIIDKDAFKVTVNGHQLELTKEYDLYLLAENKNHVMQREQIL 120
 QY 174 NHWGYNSEVTNVVDVYIYRLNKLKPYDRDKMIETVRGVGVIR 219
 DB 121 DHWGYNSEVTNVVDVYIYRLNKLKPFNKEKSIETVRGVGVIR 166

RESULT 7
 ADC96909
 ID ADC96909 standard; protein; 228 AA.

XX AC ADC96909;
 XX DT 01-JAN-2004 (first entry)
 XX DE B. faecium protein sequence SEQ ID 6536.
 XX KW Vaccine; urinary tract infection; bacteraemia; endocarditis; wound;
 XX KW abdominal-pelvic infection.
 XX OS Enterococcus faecium.
 XX PN US6583275-B1.
 XX PD 24-JUN-2003.
 XX PF 30-JUN-1998; 98US-00107532.
 XX PR 02-JUL-1997; 97US-0051571P.
 XX PR 14-MAY-1998; 98US-0085598P.
 XX PA (GENO-) GENOME THERAPEUTICS CORP.
 XX PI Doucette-Stamm LA, Bush D;
 XX WPI; 2003-799836/75.
 XX N-PSDB; ADC93255.

XX New isolated nucleic acid derived from Enterococcus faecium encoding an
 XX Enterococcus faecium polypeptide useful for detection, prevention and
 XX treatment of a pathological condition resulting from a bacterial
 XX infection.
 XX Example 1; SEQ ID NO 6536; 243pp; English.
 XX The invention relates to an isolated nucleic acid derived from
 XX Enterococcus faecium encoding an Enterococcus faecium polypeptide having

one of 10 fully defined sequences given in the (or comprising 40 sequential nucleotides chosen from any of the nucleic acids, its complement or sequences hybridising to it). Also included are a recombinant vector comprising the nucleic acid operably linked to transcription regulatory element, a cell comprising the vector and a single-stranded probe comprising the nucleic acid. The nucleic acids are chosen from 3654 disclosed sequences encoding 3654 disclosed proteins. The nucleic acids is useful for diagnosing pathological conditions resulting from *E. faecium* bacterial infection (e.g. urinary tract infection, bacteraemia, endocarditis, wounds and abdominal-pelvic infection) and for screening drugs such as agonists and antagonists. The nucleic acid is useful for recombinant production of *Candida albicans* - derived peptides or antisense polypeptides. Pharmaceutical compositions and vaccines containing the nucleic acid are useful for preventing or treating *Enterococcus faecium* infections. The present sequence represents one if the disclosed *E. faecium* proteins.

Query Match		58.3%;	Score 654;	DB 7;	Length 228;	
Best Local Similarity		56.4%;	Pred. NO. 7e-56;			
Matches 133;	Conservative	35;	Mismatches	42;	Indels	26; Gaps 4

Qy	1	MTQILLIIEDEONLAFLELELTHENYNVDTEYDQGGLDKALSHYYDLILTLMLPSING	60
Db	1	: :	
Qy	61	LEICRKIROQQSTPIIIITAKSDTYDKVAGLDYGDADYIVKPFIDIELLARIRAILRRQP	120
Db	61	: :	
Qy	121	QXDIIDVNGITIDKNAPKVTV-----NGAE-IELTKTEYDLYLLAENK	163
Db	119	----IDIEG---DKNVAQTITTYRDLTPIEKENVRWRGNEVIETLKREYELLTLTMENV	171

Qy 164 NHVQREQILNHVGYNSEVEITNVVDVYIRYLNRNKLKPYDRDKMETVRGVGVIR 219
| | : : | | | | : | | | | | | | | : | | | | | |
Dd 172 NVVLARDVLNLKVGYTEVEITNVVDVYIRYLNRNKIDVPGEESYTOTVRGTGYMR 227

RESULT 8
ABB48293
ID ABB48293 standard; protein; 236 AA.
XX
AC ABB48293;
XX
DT 05-FEB-2002 (first entry)
XX
DE Listeria monocytogenes protein #997.
XX
KW Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;
KW vitamin B12; bacterial infection; disease.
XX
OS Listeria monocytogenes.
XX
PN WO200177335-A2.
XX
PD 18-OCT-2001.
XX
PF 11-APR-2001; 2001WO-FR001118.
XX
PR 11-APR-2000; 2000FR-00004629.
XX
PA (INSP) INST PASTEUR.

Buchrieser C, Frangeul L, Couve E, Rusniok C, Fsihi H, Deboux P;
 Dausgert O, Chetouani F, Nedjari H, Glaser P, Kunst F, Cossart P;
 Daniels J, Goebel K, Krest J, Kuhn M, Ng E, Vazquez-Boland JA;
 Dominguez-Bernal G, Garrido-Garcia P, Tierrez-Martinez A, Amend A;
 Chakaraborty T, Domann E, Hain T, Berche P, Charbit A, Durant L;
 Perez-Diaz J, Baquero F, Garcia del Portillo F, Gomez-Lopez N;
 Madueno E, De Pablos B, Wehland J, Kaerst U, Entian K, Hauf J;
 Rose M, Voss H;

WPI; 2002-010914/01.

Genomic sequence for *Listeria monocytogenes*, useful e.g. for treatment and prevention of *Listeria* and related bacterial infections, and related polypeptides.

Claim 6; SEQ ID NO 998; 192bp; French.

The present invention relates to the genome sequence of *Listeria monocytogenes* EGD-e (see ABA03041). The genome sequence and fragments of it are useful for selecting probes and primers for detecting genes in *L. monocytogenes* and related organisms, and for studying genetic polymorphisms and other genomes. The present sequence is a protein encoded by the genome sequence of the present invention. Proteins expressed from the genome sequence are useful for raising specific antibodies, identification of *L. monocytogenes* and related organisms, and for biosynthesis and biodegradation, especially biosynthesis of Vitamin B12. The genome sequence and proteins encoded by it are also useful for selecting compounds that regulate gene expression and cell replication and modulate *L. monocytogenes*-related diseases. In addition, the genome sequence and proteins encoded by it are useful in pharmaceutical and vaccine compositions for the treatment or prevention of infections by *L. monocytogenes* and related organisms. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

Sequence 226 AA;

Query Match	56.0%	Score 628;	DB 5;	Length 226;
Best Local Similarity	55.1%;	Pred. No. 2.5e-53;		
Matches 124;	Conservative 39;	Mismatches 56;	Indels 6;	Gaps 1;
Qy	1	MTQILIVDEONLARFLELTHENYVNDTEYDQDGLKALSHYDYLIDLMWPSING	60	
Db	1	MNRILIVDEKUNARFIELEHQENYETAVANDCRAGLELALNEEDWAILDLMPLHNG	60	
Qy	61	LETCKIROOOSTPIIIITAKSTDYDKVAGLDVGADDIYVKPPIIBELLARIPAILRR--	118	
Db	61	VEYCRVRVQVKQTPITIMTARDSVIDRVSGLDHGADDIYVKPPAIBELLARSLRLRRVE	120	
Qy	119	---QPKQIIDVNGITIDKNAPKTVNGAEITLTKTEYDLYLLAENKNHVMQROILN	174	
Db	121	NAEQSAKQTLQYRNLIIVEKENRIVKRDSEIIDLTRREVELLTLTMENVNVITREVLN	180	
Qy	175	HWGYNSEVETNVVDYIIRLNRKLPYDRDKMIETVRGVGVIR	219	
Db	181	KWGYETVETNVVDYVYLRANKIDHPDEESYIQVRGTGYMR	225	

RESULT 9	
ADB09376	
ID	ADB09376 standard; protein; 229 AA.
XX	
XX	ADB09376;
XX	
DT	20-NOV-2003 (first entry)
XX	
DE	Alloiococcus otitis antigenic protein SEQ ID NO:3316.
XX	
KW	Alloiococcus otitidis; antigenic protein; immunogenic; immunisation;
KW	gene therapy; Gram-positive bacterium; infection.
XX	
OS	Alloiococcus otitis.

PR 18-NOV-2002; 2002US-0426742P.
 XX (AMHP) WYETH HOLDINGS CORP.
 PA Fletcher LD, Memichael JC, Russell DP, Zagursky RJ;
 PI WPI; 2003-505284/47.
 XX N-PSDB; ADB09375.
 DR New Allostococcus otitidis polynucleotides and polypeptides, useful for
 XX treating and diagnosing diseases, drug screening assays and monitoring of
 PT effects during drug clinical trials.
 XX Claim 33; SEQ ID NO 3316; 1019pp; English.
 PS
 CC The present invention describes an isolated polynucleotide (I) of
 CC Allostococcus otitidis genomic DNA, which encodes an antigenic protein.
 CC Allostococcus otitidis is a Gram-positive bacterium. Also described: (1)
 CC an isolated polypeptide that is encoded by the polynucleotide (I); (2) an
 CC expression vector comprising the novel isolated polynucleotide (I), its
 CC complement, degenerate variant or fragment; (3) a genetically engineered
 CC host cell, transfected, transformed or infected with the vector of (2);
 CC (4) an antibody specific for the polypeptide of (1); (5) an immunogenic
 CC composition comprising the polypeptide, its complement, biological
 CC equivalent or fragment, or the polynucleotide that is comprised in the
 CC expression vector; (6) a pharmaceutical composition comprising the
 CC polypeptide of (1) and a carrier; (7) a protein chip comprising an array
 CC of the polypeptides of (1), their biological equivalent or fragment; (8)
 CC immunising against Allostococcus otitidis by administering to a host the
 CC immunogenic composition; (9) detecting and/or identifying Allostococcus
 CC otitidis in the biological sample; (10) a kit comprising a container
 CC containing the novel polynucleotide, its degenerate variant or fragment,
 CC or the antibody of (4); and (11) producing a polypeptide by culturing the
 CC genetically engineered host cell under conditions suitable to produce the
 CC polypeptide from the culture. (1) can be used in gene therapy. The
 CC polynucleotides, polypeptides, antibodies and compositions of the present
 CC invention can be used for treating and diagnosing diseases, drug
 CC screening assays and monitoring of effects during drug clinical trials.
 CC The polynucleotides are useful for expressing and detecting Allostococcus
 CC otitidis. The present sequence represents an Allostococcus otitidis
 CC antigen protein from the present invention.
 XX
 SQ Sequence 229 AA;
 Query Match 52.8%; Score 592.5; DB 6; Length 229;
 Best Local Similarity 50.9%; Pred. No. 7.8e-50;
 Matches 116; Conservative 46; Mismatches 57; Indels 9; Gaps 2;
 QY 1 MTOILIVEDONLARFLELELTHENYNVDTEYDQDGLDKALS-HYYDLIILDLMLPSIN 59
 DB 1 MKKVLIEDENKALRFVELELKHGEGFAPTVKFGDSGLQALNGEDWDVLLDMLPELN 60
 QY 60 GLEICRKIROQOSTPIIIITAKSDTYDKVAGLDYGDADDYIVKPFDELLARILRR- 118
 DB 61 GIDVARRVRQVSEPIIMMTARDSDVIDRVSGFDHGADDYLVKPPAIEELLARALPRRI 120
 QY 119 -----OPQKDIDVNGITIDKNAFKVTNGABIELTKTEYDOLLYLAENKNHVMQREQ 171
 DB 121 DIESSIQEKQKILEYDLRVLEKENRIVRRDDEVIETLTREYELLVELMENVDVLSREK 180
 QY 172 ILNHVMGYNSEVTNVVDVYIRYLRLNKLKPYDRDKMIETVRGVGYVIR 219
 DB 181 LLNEVWGESDIENVVDVYIRYLRLNKLIDRPTGESYIQTVRGTGYVMR 228
 RESULT 10
 ID ABB54940
 XX ABB54940 standard; protein; 230 AA.
 AC ABB54940;
 XX
 XX 29-AUG-2003 (revised)
 DT 16-MAY-2002 (first entry)

XX Lactococcus lactis protein lrrA.
 DE
 XX Biosynthesis; biodegradation; lactic bacterium; yogurt; cheese.
 KW
 XX Lactococcus lactis; IL1403.
 OS
 XX PR2807446-A1.
 PN
 XX 12-OCT-2001.
 PD
 XX 11-APR-2000; 2000FR-00004630.
 PF
 XX 11-APR-2000; 2000FR-00004630.
 PR
 XX (INRG) INRA INST NAT RECH AGRONOMIQUE.
 PA
 XX Bolotine A, Sorokine A, Renault P, Ehrlich SD;
 PI
 XX WPI; 2002-043418/06.
 DR
 XX New nucleotide sequence useful in the identification of Lactococcus
 PT lactis and related species.
 PT
 XX Claim 6; SEQ ID NO 1642; 2504pp; French.
 PS
 XX The present invention is related to a Lactococcus lactis nucleotide
 CC sequence (ABA90521) and related proteins (ABB53300-ABB55621). The nucleic
 CC acid sequence is useful in the detection and/or amplification of nucleic
 CC acid sequence, particularly to identify Lactococcus lactis or related
 CC species. The proteins of the invention are useful for the biosynthesis or
 CC biodegradation of a composition of interest. The invention helps research
 CC in lactic bacteria, particularly useful in the production of yogurt and
 CC cheese. Note: The sequence data for this patent is based on equivalent
 CC patent WO20017734 (published 18-OCT-2001) which is available in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences. (Updated on 29-AUG-2003 to
 CC standardise OS field)
 XX
 SQ Sequence 230 AA;
 Query Match 52.0%; Score 583; DB 5; Length 230;
 Best Local Similarity 51.6%; Pred. No. 6.7e-49;
 Matches 115; Conservative 40; Mismatches 62; Indels 6; Gaps 1;
 QY 3 QILIVEDONLARFLELELTHENYNVDTEYDQDGLDKALSHYDYLILDLMLPSINGLE 62
 DB 5 KILIIEDENKALRFVLSLEHEGYATEIKDNGRSGLEEATSKDYDILDLMLPELDGFE 64
 QY 63 ICRKIROQOSTPIIIITAKSDTYDKVAGLDYGDADDYIVKPFDELLARILRRPOK 122
 DB 65 VARRLEKEDTPIIMMTARDSDMDRVAGLDIGADDYITKPPAIEELLARVRAFFREEHG 124
 QY 123 DIID-----VNGITIDKNAFKVTNGABIELTKTEYDOLLYLAENKNHVMQREQINHV 176
 DB 125 HAVERAENTSPRDLVDIKTNRTVHRGKVIDLTRREYDILLTLMLQNVGVVTRHLVSQV 184
 QY 177 WGNSEVTNVVDVYIRYLRLNKLKPYDRDKMIETVRGVGYVIR 219
 DB 185 WGYEETETNVVDVYIRYLRLNKLIDVEGQDSYIQTVRGLGYVMR 227
 RESULT 11
 ID ABB27835
 XX ABB27835 standard; protein; 229 AA.
 AC ABB27835;
 XX
 XX 02-JUL-2002 (first entry)
 DT
 XX Streptococcus polypeptide SEQ ID NO 4846.
 DE
 XX Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
 KW

KW group A streptococcus; Streptococcus pyogenes; antibacterial;
 XX antiinflammatory; infection; vaccine; meningitis; gene therapy.

OS Streptococcus agalactiae.

PN WO200234771-A2.

PD 02-MAY-2002.

XX 29-OCT-2001; 2001WO-GB004789.

XX 27-OCT-2000; 2000GB-00026333.

PR 24-NOV-2000; 2000GB-00028727.

PR 07-MAR-2001; 2001GB-00005640.

XX (CHIR-) CHIRON SPA.

PA (GENO-) INST GENOMIC RES.

XX Telford J, Masignani V, Margarit Y Rosl, Grandi G, Fraser C;

PI Tettelin H;

XX WPI; 2002-352536/38.

DR N-PSDB; ABN68466.

XX New Streptococcus protein for the treatment or prevention of infection or

PT disease caused by Streptococcus bacteria, such as meningitis, and for

PT detecting a compound that binds to the protein.

XX Claim 1; Page 3647; 4525pp; English.

XX The invention relates to a protein (ABP25413-ABP30895) from group B
 CC streptococcus/GAS (Streptococcus agalactiae) or group A streptococcus/GAS
 CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in
 CC the specification. The proteins have antibacterial and antiinflammatory
 CC activity. (I), nucleic acids encoding (I), ABN66044-ABN71526 and
 CC antibodies that bind (I) are used in the manufacture of medicaments for
 CC the treatment or prevention of infection or disease caused by

CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.

CC Nucleic acids encoding (I) are used to detect Streptococcus in a

CC biological sample. (I) is used to determine whether a compound binds to

CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be

CC used as a vaccine or diagnostic composition. The disease caused by

CC Streptococcus that is prevented or treated may be meningitis. Nucleic

CC acid encoding (I) may be used to recombinantly produce (I) and may be

CC used in gene therapy. Antibodies to (I) are used for affinity

CC chromatography, immunoassays, and distinguishing/identifying

CC Streptococcus proteins

XX Sequence 229 AA;

Query Match 51.2%; Score 574.5; DB 5; Length 229;

Best Local Similarity 53.3%; Pred. No. 4.6e-48;

Matches 121; Conservative 38; Mismatches 55; Indels 13; Gaps 3;

QY 3 QILIVEDQNLFLELELTHENVNVDYDQGLDKALSHYVDLIIIDMLPSINGLE 62

DB 4 KILIIEDENLARFVSLLEHGVVETNREGDGTALEKDFDILLDLMPMDGFE 63

QY 63 ICRKIROQOSTPIIIITAKSDTYDKVAGLDYDADDYVKKPDIPELLARIRAILRRQOK 122

DB 64 ITRRLQAEKTTYIMMTARDSYMDIVAGLDRGADDYVKKPFAIBELLARVRAIFRRQIE 123

QY 123 DIIDVNGITIDKNAFK-VTVN-----GAEIELTKTEYDLYLLAENKHNVMQREOI 172

DB 124 TTKTEKG---DSGSFRDISLNTHNSAMRGDEISLTKEFDLLNVLMTNVMNRVMTREEL 180

QY 173 LNHVNGYNEVETNVVDYVIRYLNRKLKPYDRDKMETVGRGVYIR 219

DB 181 LEHWKYDVAETNVVDYVIRYLGRKIDIPGRESYIQTVRGNGYVIR 227

RESULT 12

ABP27042

ABP27042 standard; protein; 228 AA.

AC ABP27042;

XX 02-JUL-2002 (first entry)

XX Streptococcus polypeptide SEQ ID NO 3260.

XX Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
 KW group A streptococcus; Streptococcus pyogenes; antibacterial;
 KW antiinflammatory; infection; vaccine; meningitis; gene therapy.

XX Streptococcus pyogenes.

OS WO200234771-A2.

PN 02-MAY-2002.

XX 29-OCT-2001; 2001WO-GB004789.

XX 27-OCT-2000; 2000GB-00026333.

PR 24-NOV-2000; 2000GB-00028727.

PR 07-MAR-2001; 2001GB-00005640.

XX (CHIR-) CHIRON SPA.

PA (GENO-) INST GENOMIC RES.

XX Telford J, Masignani V, Margarit Y Rosl, Grandi G, Fraser C;

PI Tettelin H;

XX WPI; 2002-352536/38.

DR N-PSDB; ABN67673.

XX New Streptococcus protein for the treatment or prevention of infection or
 CC disease caused by Streptococcus bacteria, such as meningitis, and for
 CC detecting a compound that binds to the protein.

XX Claim 1; Page 3480; 4525pp; English.

XX The invention relates to a protein (ABP25413-ABP30895) from group B
 CC streptococcus/GAS (Streptococcus agalactiae) or group A streptococcus/GAS
 CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in
 CC the specification. The proteins have antibacterial and antiinflammatory
 CC activity. (I), nucleic acids encoding (I), ABN66044-ABN71526 and
 CC antibodies that bind (I) are used in the manufacture of medicaments for
 CC the treatment or prevention of infection or disease caused by

CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.

CC Nucleic acids encoding (I) are used to detect Streptococcus in a

CC biological sample. (I) is used to determine whether a compound binds to

CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be

CC used as a vaccine or diagnostic composition. The disease caused by

CC Streptococcus that is prevented or treated may be meningitis. Nucleic

CC acid encoding (I) may be used to recombinantly produce (I) and may be

CC used in gene therapy. Antibodies to (I) are used for affinity

CC chromatography, immunoassays, and distinguishing/identifying

CC Streptococcus proteins

XX Sequence 228 AA;

Query Match 51.2%; Score 574; DB 5; Length 228;

Best Local Similarity 52.9%; Pred. No. 5.1e-48;

Matches 119; Conservative 38; Mismatches 58; Indels 10; Gaps 2;

QY 3 QILIVEDQNLFLELELTHENVNVDYDQGLDKALSHYVDLIIIDMLPSINGLE 62

DB 4 KILIIEDENLARFVSLLEHGVVETNREGDGTALEKDFDILLDLMPMDGFE 63

QY 63 ICRKIROQOSTPIIIITAKSDTYDKVAGLDYDADDYVKKPDIPELLARIRAILRRQ--- 119

DB 64 VTRRLQAEKTTYIMMTARDSYMDIVAGLDRGADDYVKKPFAIBELLARVRAIFRRQIE 123

QY 120 -----PQKDIIDVNGITIDKNAFKVTVNGABIELTKTEYDLYLLAENKHNVMQREQLN 174

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 4, 2004, 18:00:48 ; Search time 16.997 Seconds
(without alignments)
1239.391 Million cell updates/sec

Title: US-09-006-627-4
Perfect score: 1122
Sequence: 1 MTQLIVEDEQNLFARFLE.....KPYDRDKMIETVRGVGVYIR 219

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues
Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	628	56.0	226	2 AE1609	two-component resp
2	628	56.0	226	2 A11246	two-component resp
3	583	52.0	230	2 B86824	two-component syst
4	539	48.0	222	2 D3696	two-component resp
5	527	47.0	228	2 C59859	two-component resp
6	526.5	46.9	228	2 AG1955	two-component resp
7	490.5	43.7	232	2 G97109	response regulator
8	488.5	43.5	225	2 H75480	response regulator
9	482.5	43.0	224	1 S76529	hypothetical prote
10	481	42.9	238	2 D33847	two-component resp
11	479	42.7	238	2 AD1690	two-component resp
12	476	42.4	238	2 AD1318	two-component resp
13	475	42.3	239	2 E84044	two-component resp
14	472	42.1	240	1 G69691	two-component resp
15	463.5	41.3	257	2 D70624	probable two-compo
16	458	40.8	230	2 C56939	response regulator
17	451.5	40.2	228	2 A97296	response regulator
18	449	40.0	235	2 B97660	response regulator
19	447	39.8	233	2 T45359	two-component regu
20	447	39.8	236	2 D70783	probable two compo
21	446.5	39.8	248	2 AC0630	response regulator
22	445.5	39.7	225	2 H72382	response regulator
23	445	39.7	230	2 A70821	hypothetical prote
24	445	39.7	235	2 S65967	two-component resp
25	444.5	39.6	228	2 H70592	probable mtrA prot
26	442.5	39.4	228	2 F87005	probable two-compo
27	442	39.4	234	2 D55142	DNA-binding respon
28	442	39.4	238	2 T31680	baCR protein - Bac
29	442	39.4	247	2 D72228	response regulator

30	442	39.4	251	2 B98010	response regulator
31	440.5	39.3	221	2 C96980	response regulator
32	440.5	39.3	226	1 S70164	copper resistance
33	439	39.1	241	2 D89928	hypothetical prote
34	438.5	39.1	227	2 G89782	two-component resp
35	438.5	39.1	228	2 B86750	two-component syst
36	436	38.9	233	2 E86674	two-component resp
37	436	38.9	236	2 C94153	two-component resp
38	435.5	38.8	224	2 H83330	probable two-compo
39	432.5	38.5	234	2 B89953	hypothetical prote
40	432.5	38.5	237	2 AD1472	two-component resp
41	432.5	38.5	237	2 AH1110	two-component resp
42	432.5	38.5	242	2 AG2283	two-component resp
43	432	38.5	230	2 E87285	phosphate regulon
44	431	38.4	236	2 AE1387	two-component resp
45	430.5	38.4	239	2 E64961	probable transcript

ALIGNMENTS

RESULT 1
AE1609
two-component response regulator lisR [imported] - Listeria innocua (strain Clip11262)
C;Species: Listeria innocua
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
C;Accession: AE1609
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A;Authors: Krest, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A.; Title: Comparative genomics of Listeria species.
A;Reference number: AB1077; MUID:21537279; PMID:11679669
A;Accession: AE1609
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-226 <GLA>
A;Cross-references: GB:AL592022; PIDN:CAC96645.1; PID:gl6413887; GSPDB:GN00178
A;Experimental source: strain Clip11262
C;Genetics:
A;Gene: lisR
C;Superfamily: ompR protein; response regulator homology

Query Match	56.0%	Score 628	DB 2	Length 226
Best Local Similarity	55.1%	Pred. No. 3.9e-40		
Matches 124	Conservative 39	Mismatches 56	Indels 6	Gaps 1
Qy	1	MTQLIVEDEQNLFARFLELELTHENYNVDTEYDQGLDKALSHYDILIIIDMLPSING 60		
Db	1	MNRILIVEDEKNLARFIELELOHENYETAVANDGRAGLEALNEEDAILLDMLPHUNG 60		
Qy	61	LEICRKTRQOOSTPIIIITAKSDTYKVAGLDVGADYIVKVPFIEELLARAILRR-- 118		
Db	61	VEVCRVRVQKQTPFIIMITARDSVIDRVSGLDGADYIVKVPFIEELLARLRLRRVE 120		
Qy	119	----QPQKDIIDVNGITIDKNAPKVTNGAEIETKTEYDILLYLLAENKHVMOREQILN 174		
Db	121	NAEOSAKQTLQVRNKLIVEKENRIVKDEBIIDTKREYELLTLTMENVILTVREVLN 180		
Qy	175	HVMGYNSEVTNVVDVYIRLRNKLKPYDRDKMIETVRGVGVYIR 219		
Db	181	KWGYEYETNVVDVYVIRLRNKLIDHPDESYIQTVRGTGYVNR 225		

RESULT 2
A11246
two-component response regulator lisR [imported] - Listeria monocytogenes (strain EGD-e)
C;Species: Listeria monocytogenes
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
C;Accession: A11246
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker

.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A;Authors: Krefte, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlant,
A;Title: Comparative genomics of *Listeria species*.
A;Reference number: AB1077; MUID:21537279; PMID:11679669
A;Accession: A11246
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-226 <GLA>
A;Cross-references: GB:NC_003210; PIDN:CAC99455.1; PID:g16410793; GSPDB:GN00177
A;Experimental source: strain EGD-e
C;Genetics:
A;Gene: *lirA*
C;Superfamily: ompR protein; response regulator homolog

Query Match 56.0%; Score 628; DB 2; Length 226;
Best Local Similarity 55.1%; Pred. No. 3.9e-40;
Matches 124; Conservative 39; Mismatches 56; Indels 6; Gaps 1;
QY 1 MTQILIVDEQNARLFLELELTHENNVNVDTEYDQDGLDKALSHYVDLIILDLMLPSNG 60
DB 1 MNRILIVDEKNLARFLELELTHENNVNVDTEYDQDGLDKALSHYVDLIILDLMLPSNG 60
QY 61 LEICRKIROQOSTPIIIITAKSDTYKVAGLDYGDYIVKPFDEIELLARAILR-- 118
DB 61 VEVCRVRQVKOTPIIMTARDSVIDRSGLDGADYIVKPFDEIELLARAILR-- 120
QY 119 ----QPOKDIIDNGITIDKNAPKVTNGAEIETKTETDLYLLAENKNHVMQREQLN 174
DB 121 NAOQSAQTTLQVRLNIVKRNKIDHPDESYIQTVRGTYGVYR 219
QY 175 HVGYNSEVTNVVDYVIRYLRNKLKPYDRDKMIETVRGVGYVIR 219
DB 181 KWGYETEVTNVVDYVIRYLRNKLKIDHPDESYIQTVRGTYGVYR 225

RESULT 3
B86824
two-component system regulator *lirA* [imported] - Lactococcus lactis subsp. lactis (strain
C;Species: Lactococcus lactis subsp. lactis
C;Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001
C;Accession: B86824
R;Bolotin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarne, K.; Weissenbach, J.; Ehrlich,
Genome Res. 11, 731-753, 2001
A;Title: The complete genome sequence of the lactic acid bacterium *Lactococcus lactis* s
A;Reference number: A86625; MUID:21235186; PMID:11337471
A;Accession: B86824
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-230 <STO>
A;Cross-references: GB:AE005176; PID:g12724600; PIDN:AAK05692.1; GSPDB:GN00146
A;Experimental source: strain IL1403
C;Genetics:
A;Gene: *lirA*
C;Superfamily: ompR protein; response regulator homolog

Query Match 52.0%; Score 583; DB 2; Length 230;
Best Local Similarity 51.6%; Pred. No. 9.5e-37;
Matches 115; Conservative 40; Mismatches 52; Indels 6; Gaps 1;
QY 3 QILIVDEQNARLFLELELTHENNVNVDTEYDQDGLDKALSHYVDLIILDLMLPSNGLE 62
DB 5 KILIIIEKNLARFVLELEHEGYATEIKNGRSGLEATSKDYDILLOLMLPELDGFE 64
QY 63 ICRKIROQOSTPIIIITAKSDTYKVAGLDYGDYIVKPFDEIELLARAILRQPOK 122
DB 65 VARRLRKEKTPPIIMTARDSMDRVAGLDIGADYITKPFDEIELLARVAFRRHEHG 124
QY 123 DIID-----VNGITIDKNAPKVTNGAEIETKTETDLYLLAENKNHVMQREQLNHV 176
DB 125 HAVERAENTSFDLVIDKTNRTVHRGKKVIDLTREYDILLTLNQNVGDVVTRHLVSQV 184

QY 177 WGVNSEVTNVVDYVIRYLRNKLKPYDRDKMIETVRGVGYVIR 219
DB 185 WGYEGTETNVVDYVIRYLRNKLKIDVEGQDSYIQTVRGLGYVMR 227

RESULT 4

D83696

two-component response regulator BH0372 [imported] - *Bacillus halodurans* (strain C-125)
C;Species: *Bacillus halodurans*
C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C;Accession: D83696
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hiran
Nucleic Acids Res. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and
A;Reference number: A83650; MUID:20512582; PMID:11058132
A;Accession: D83696
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-222 <STO>
A;Cross-references: GB:AP001508; GB:BA000004; NID:g10172890; PIDN:BAB04091.1; GSPDB:GN001
A;Experimental source: strain C-125
C;Genetics:
A;Gene: BH0372
C;Superfamily: ompR protein; response regulator homolog

Query Match 48.0%; Score 539; DB 2; Length 222;
Best Local Similarity 48.2%; Pred. No. 1.8e-33;
Matches 108; Conservative 45; Mismatches 63; Indels 8; Gaps 3;
QY 1 MTQILIVDEQNARLFLELELTHENNVNVDTEYDQDGLDKALSHYVDLIILDLMLPSNG 60
DB 1 MARILIIIEKRIARFLELELTHENNVNVDTEYDQDGLDKALSHYVDLIILDLMLPSNG 60

QY 61 LEICRKIROQOSTPIIIITAKSDTYKVAGLDYGDYIVKPFDEIELLARAILR-- 117
DB 61 LEVLRIRWTDPTPIILLTARNIPDKVSGLDLGANDYITKPFDEIELLARVACLRTV 120

QY 118 --RQPOKDIIDNGITIDKNAPKVTNGAEIETKTETDLYLLAENKNHVMQREQLNH 175
DB 121 QTRRVEDTLMPQELTINEKTRDQVQGNETIETPKFELLVFFIKNGQVLSREQILT 180

QY 176 WGVNSEVTNVVDYVIRYLRNKLKPYDRDKMIETVRGVGYVIR 219
DB 181 WGVFDYGTNVVDYVIRYLRNKLKPYDRDKMIETVRGVGYRDK 221

RESULT 5

C69859

two-component response regulator [YkoH] homolog ykoG - *Bacillus subtilis*

C;Species: *Bacillus subtilis*

C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 20-Jun-2000

C;Accession: C69859

R;Kunst, F.; Ogawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho
A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997

A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gall
A.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.;
Koester, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel
Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle,
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanlon,
A;Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seror,
akeuchi, M.; Tanakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.
A;Authors: Yoshikawa, H.F.; Zumstein, E.; Yoeikawa, H.; Danchin, A.

A;Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.

A;Reference number: A69580; MUID:98044033; PMID:9384377

A;Accession: C69859

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-228 <UN>

A;Cross-references: GB:Z99110; GB:AL009126; NID:92633472; PIDN:CAB13182.1; PID:g2633679
A;Experimental source: strain 168
C;Genetics:
A;Gene: ykoG
C;Superfamily: ompR protein; response regulator homology
C;Keywords: phosphoprotein
F;6-115/Domain: response regulator homology <RRH>
F;54/Binding site: phosphate (Asp) (covalent) #status predicted

Query Match 47.0%; Score 527; DB 2; Length 228;
Best Local Similarity 48.4%; Pred. No. 1.5e-32;
Matches 108; Conservative 50; Mismatches 5; Indels 8; Gaps 5;

QY 4 ILIIVDEQNARFLELELTHENYNVDTEYDQDGLKALSHYDYLIIIDLMPLSINGLEI 63
DB 6 ILIIVDEEKIARVLQLELEYEGSVTIKINGTEGLDAAAGGYSIVLVDVMLPGLSGLV 65

QY 64 CRKIRQQQS--TPIIITAKSDTYDKVAGLDYGADYIVKFPDIEELLARIRAILRRQPK 122
DB 66 LRLRLKTSQTPVILLTARDSPDKVTGLDYGANDYVTKPFEIEELLARIRAILRQNGTK 125

QY 123 --DI---IDVNGITDKNAFKVTNGABIELTKTEYDLYLLAENKNHVMQREQLNHVW 177
DB 126 TEDICTFTYDRLRNEKTRVGRGKEVELTPREFDLVYMLKHPQQVLTREQLSSVW 185

QY 178 GYNSEVTNVVDVYIRYLNRNKLK-PYDRDKMIETVRGVGVYIR 219
DB 186 GPDYIGDNTVVDVYIRYIRKLDYDYEK-QLIHTIRGVGYAIK 227

RESULT 6
AG1955
C;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
C;Accession: AG1955
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, Nakazaki, N.; Shimoto, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S. DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AG1955
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-228 <KUR>
A;Cross-references: GB:BA000019; PIDN:BA073151.1; PID:gl7130541; GSPDB:GN00179
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: alr1194
C;Superfamily: ompR protein; response regulator homology

Query Match 46.9%; Score 526.5; DB 2; Length 228;
Best Local Similarity 48.4%; Pred. No. 1.6e-32;
Matches 107; Conservative 41; Mismatches 70; Indels 3; Gaps 2;

QY 2 TOILIVEQNARFLELELTHENYNVDTEYDQDGLKALSHYDYLIIIDLMPLSINGL 61
DB 3 THILLIVEDEVKLARFVLELSEGVKSVAGHDGTLTLARESTPDVLAVLDWMLPGLSGL 62

QY 62 EICRKIRQQ-OSTPIIIITAKSDTYDKVAGLDYGADYIVKFPDIEELLARIRAILRR-- 118
DB 63 EICRLRATGNSIPVILLTARDSEVSDRVAGLDAGADDYVVKPFSIEELLARIRAILRRRTQ 122

QY 119 QPKDIIIDVNGITDKNAFKVTNGABIELTKTEYDLYLLAENKNHVMQREQLNHVW 178
DB 123 ETDEDLQFDSLNRRTREVFRGNRAVELTAKEFDLLEYLLSVYPRQVFTTRDQILEKVG 182

QY 179 YNSEVETNVVDVYIRYLNRNKLKPYDRDKMIETVRGVGVYIR 219
DB 183 YDFMGDSNIIIEVYIRYLKLEENNEKRLVHTVRGVGYALR 223

RESULT 7
G97109
response regulator (Chey-like receiver domain and DNA-binding HTH domain) [imported] - C
C;Species: Clostridium acetobutylicum
C;Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 30-Sep-2001
C;Accession: G97109
R;Nolling, J.; Breton, G.; Omeichenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee, J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clostridium acetobutylicum ATCC824
A;Reference number: A96900; MUID:21359325; PMID:21359325
A;Accession: G97109
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-232 <KUR>
A;Cross-references: GB:AF001437; PIDN:AAK79666.1; PID:gl5024664; GSPDB:GN00168
A;Experimental source: Clostridium acetobutylicum ATCC824
C;Genetics:
A;Gene: CAC1700
C;Superfamily: ompR protein; response regulator homology

Query Match 43.7%; Score 490.5; DB 2; Length 232;
Best Local Similarity 44.1%; Pred. No. 8.5e-30;
Matches 98; Conservative 50; Mismatches 65; Indels 9; Gaps 4;

QY 3 QILIVDEQNARFLELELTHENYNVDTEYDQDGLKALSHYDYLIIIDLMPLSINGLE 62
DB 5 KILVVDDEHIVKLIKFNLENNYKVITAADGGEALKEKAKGEVQVLVLDLMLPMDGYD 64

QY 63 ICRKIRQQQS---TPIIITAKSDTYDKVAGLDYGADYIVKFPDIEELLARIRAILRRQ 119
DB 65 VCRBIRDQSDISNPNVIMITAKGEELDKILGALGADYITKPFSEVRLVARVAVLRR- 123

QY 120 PQKDIID----VNGITDKNAFKVTNGABIELTKTEYDLYLLAENKNHVMQREQLNH 175
DB 124 TKVDYIDTKFKGNIQIDFQRHNVTKSGEKVELTKFELQLVLIKNKGRVWTRDFLLDK 183

QY 176 VWGYNSEVTNVVDVYIRYLNRNKLKPYDRDKMIETVRGVGVY 216
DB 184 IWGEYIEGRTVDVHVRLRQKIEDDDKPKYIETIRGIGY 225

RESULT 8
H75480
C;Species: Deinococcus radiodurans (strain R1)
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000
C;Accession: H75480
R;White, O.; Eise, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.; S.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma, S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A;Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A;Reference number: A75250; MUID:20036896; PMID:10567266
A;Accession: H75480
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-225 <WHI>
A;Cross-references: GB:AF001930; GB:AE000513; NID:g6458450; PIDN:AAF10321.1; PID:g645845
A;Experimental source: strain R1
C;Genetics:
A;Gene: DR0743
A;Map position: 1
C;Superfamily: ompR protein; response regulator homology

Query Match 43.5%; Score 488.5; DB 2; Length 225;
Best Local Similarity 48.4%; Pred. No. 1.2e-29;
Matches 106; Conservative 37; Mismatches 71; Indels 5; Gaps 3;

QY 4 ILIIVDEQNARFLELELTHENYNVDTEYDQDGLKALSHYDYLIIIDLMPLSINGLEI 63
DB 7 VLVIIEKDIARFIELEAAEGYATEVAFDVGTVGLSKFREYNPOLVLDLMLPVLGLV 66

C>Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
C/Accession: AD1318
R/Glaeser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A/Authors: Kretz, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Maok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A./Title: Comparative Genomics of *Listeria* species.
A/Reference number: AB1077; MUID:21537279; PMID:11679669
A/Accession: AD1318
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-238 <GLA>
A/Cross-references: GB:NC_003210; PIDN:CAD00026.1; PID:g16411401; GSPDB:GN00177
A/Experimental source: strain EGD-e
C/Genetics:
A/Gene: read
C/Superfamily: ompR protein; response regulator homology

Query Match 42.4%; Score 476; DB 2; Length 238;
Best Local Similarity 41.2%; Pred. No. 1.1e-28;
Matches 93; Conservative 49; Mismatches 72; Indels 12; Gaps 2;

Qy 3 QILIVEDQNLARFLELELTHENVNVDYDQGDGLKALSHYDYLILDLMLPSINGLE 62
Db 6 RVLVDDDEIRRLKMYLERENYRIEASDQDQALSALANNVYEVILLDMLMPKOGIE 65

Qy 63 ICRKIRQOQSPILIIITAKSDYDKVAGLDYGADYIVKPFPIEELLARIALLRQPK 122
Db 66 VCRLREFKSPVVMVLTAKGEANRVQGFVGADYIVKPFSPREVLRVAVLRRAKQS 125

Qy 123 -----DIIDVNGITIDKNAPKVTWNGAEIELTKTEYDLYLLAENKNHVMQREQI 172
Db 126 SEESAGTTPGDIILFPPLKIDNEARHVIDGKEIGLTPKEVDLYLLAKSPKVDRESL 185

Qy 173 LNHVWYNSVETNVVDYVIRLNRKLPYDRD--KMETVRGVGY 216
Db 186 LKEVWRYEFGDLRTIDTHVKRLREKLHDVSEDAARMIVTWGLGY 231

RESULT 13
E84044
Two-component response regulator involved in phosphate regulation phop [imported] - Bacillus halodurans
C/Species: Bacillus halodurans
C/Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C/Accession: E84044
R/Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A/Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and
A/Reference number: AB3650; MUID:20512582; PMID:11058132
A/Accession: E84044
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-239 <STO>
A/Cross-references: GB:AP001517; GB:BA000004; NID:g10175500; PIDN:BA06876.1; GSPDB:GN00
A/Experimental source: strain C-125
C/Genetics:
A/Gene: phop
C/Superfamily: ompR protein; response regulator homology

Query Match 42.3%; Score 475; DB 2; Length 239;
Best Local Similarity 41.7%; Pred. No. 1.3e-28;
Matches 96; Conservative 50; Mismatches 70; Indels 14; Gaps 4;

Qy 1 MTQ-ILIVEDQNLARFLELELTHENVNVDYDQGDGLKALSHYDYLILDLMLPSIN 59
Db 1 MTQRLVDDDESIIVTLQFNLSQSGYEVVMTAMDGASGLQAKTQTFDILDLMLPEMD 60

Qy 60 GLEICRKIRQOQ--SPILIIITAKSDYDKVAGLDYGADYIVKPFPIEELLARIALLR 118
Db 61 GLDVCKQRQRKQVMTPIILMLTAKODEFDKVLGLGELGADYVMTKQFSPREVVRARILRR 120

Qy 119 -----QPKQDIDVNGITIDKNAPKVTWNGAEIELTKTEYDLYLLAENKNHVM 167
Db 121 VGHAAQNHNEDNSRSTSLTSGFNVEIYDPNVEYLKQGPQLELTPKEFELLYLANHKGRLV 180

Qy 168 QREQILANHVWYNSVETNVVDYVIRLNRKLPYDRDKM-IETVRGVGY 216
Db 181 TRDQLLANVWYNEFVGDTTRIVDHSVHLSHLEKIEPTKPKPIYKIRGLGY 230

RESULT 14
G69691
Two-component response regulator involved in aerobic and anaer resd - *Bacillus subtilis*
C/Species: *Bacillus subtilis*
C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 16-Jun-2000
C/Accession: G69691; S45559
R/Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter, C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho, A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.; Nature 390, 249-256, 1997
A/Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler, Iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.; Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois, A.; Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel, Y. M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle, Rieger, M.; Rivoita, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon, A.; Authors: Schleich, S.; Schroeter, R.; Scoffone, P.; Sekiguchi, J.; Sekowska, A.; Seror, akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpetra, P.; Tognoni, J.; Uchiyama, T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.; Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A/Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.
A/Reference number: A69580; MUID:98044033; PMID:9384377
A/Accession: G69691
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-240 <KUN>
A/Cross-references: GB:Z99116; GB:AL009126; NID:g2634723; PIDN:CAB14244.1; PID:g2634747
A/Experimental source: strain 168
R/Sorokin, A.; Zumstein, E.; Azevedo, V.; Ehrlich, S.D.; Serror, P.
submitted to the EMBL Data Library, November 1993
A/Reference number: S45533
A/Accession: S45559
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-240 <SOR>
A/Cross-references: EMBL:L05228; NID:g410114; PIDN:AAA67497.1; PID:g410141
C/Genetics:
A/Gene: resd
C/Superfamily: ompR protein; response regulator homology
C/Keywords: phosphoprotein
F:9-117/Domain: response regulator homology <RRH>
F:57/Binding site: phosphate (Asp) (covalent) #status predicted

Query Match 42.1%; Score 472; DB 1; Length 240;
Best Local Similarity 41.2%; Pred. No. 2.2e-28;
Matches 94; Conservative 51; Mismatches 69; Indels 14; Gaps 3;

Qy 2 TQILIVEDQNLARFLELELTHENVNVDYDQGDGLKALSHYDYLILDLMLPSINGL 61
Db 7 TKILVDDDEARIRRLRLMYLERENYRIEASDQDQALSALANNVYEVILLDMLMPGTDGI 66

Qy 62 EICRKIRQOQSTPIIIITAKSDYDKVAGLDYGADYIVKPFPIEELLARIALLR--- 118
Db 67 EVCQRIREKATPIIMLTAKGEANRVQGFENGTDYIVKPFSPREVLRVAVLRRAASQ 126

Qy 119 -----QPKQDIDVNGITIDKNAPKVTWNGAEIELTKTEYDLYLLAENKNHVMQRE 171
Db 127 TSVFNANTPTKNVLFVSHLSIDHDAHRVADGTAEVSLTPKVYELLYFLAKTPDKVYDREK 186

Qy 172 ILNHVWYNSVETNVVDYVIRLNRKLPYDRDKM-IETVRGVGY 216
Db 187 LLKEVWRYEFGDLRTIDTHVKRLREKLNKLVSP-EAAKIVTWGVGY 233

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 4, 2004, 17:51:42 ; Search time 9.80597 Seconds
(without alignments)
1162.900 Million cell updates/sec

Title: US-09-006-627-4

Perfect score: 1122

Sequence: 1 MTQILIVEQNLARFLE.....KPYDRDKMIETVRGVGVIR 219

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	527	47.0	228	1 YKOG_BACSU	O34903 bacillus su
2	472	42.1	240	1 RESD_BACSU	P35163 bacillus su
3	449	40.0	225	1 QCCR_ALCEU	Q44006 alcaligenes
4	448.5	40.0	226	1 PCOR_ECOLI	Q47456 escherichia
5	447	39.8	233	1 PRRA_MYCLE	Q50136 mycobacteri
6	447	39.8	233	1 PRRA_MYCTU	O10531 mycobacteri
7	445	39.7	235	1 YVCF_BACSU	P37478 bacillus su
8	444.5	39.6	225	1 MTRA_MYCPA	Q93cb8 mycobacteri
9	444.5	39.6	225	1 MTRA_MYCTU	Q50447 mycobacteri
10	442.5	39.4	225	1 MTRA_MYCLE	Q9ccj2 mycobacteri
11	438.5	39.1	227	1 YCLJ_BACSU	P94413 bacillus su
12	437	38.9	229	1 IRLR_BURPS	O31395 burkholderi
13	430.5	38.4	223	1 IRLP_ECOLI	P76340 escherichia
14	428	38.1	240	1 PHOP_BACSU	P13792 bacillus su
15	428	38.1	240	1 YC27_PORAE	P28835 porphyridiu
16	425	37.9	230	1 VANR_ENTFC	Q06239 enterococcu
17	424.5	37.8	224	1 DLTR_STR3A	Q8dxq7 streptococc
18	423	37.7	239	1 YC27_CYAPA	P48259 cyanophora
19	422.5	37.7	225	1 AQI1_STRCO	Q04942 streptomyce
20	420	37.4	243	1 YC27_PORPU	P51358 porphyra pu
21	419	37.3	254	1 YC27_GUITH	Q78428 guillardia
22	417	37.2	228	1 RGX3_MYCSM	Q9f868 mycobacteri
23	416	37.1	227	1 COPR_PSESM	Q02540 pseudomonas
24	410.5	36.6	227	1 RGX3_MYCHO	O07130 mycobacteri
25	410.5	36.6	227	1 RGX3_MYCTU	Q11156 mycobacteri
26	410	36.5	249	1 YC27_CYACA	Q951q4 cyanidium c
27	406	36.2	229	1 CUSR_ECOLI	P77380 escherichia
28	406	36.2	229	1 PHOB_ECOLI	P08402 escherichia
29	406	36.2	229	1 YC27_SHIDY	P45606 shigella dy
30	406	36.2	255	1 YC27_GALSD	P28257 galdieria s
31	404	36.0	632	1 RCAC_FREDI	Q01473 fremyella d
32	399	35.6	229	1 PHOB_KLEPN	P45605 klebsiella
33	398	35.5	229	1 PHOB_SHIFL	P45607 shigella fl

RESULT 1

YKOG_BACSU
ID YKOG_BACSU STANDARD; PRT; 228 AA.
AC O34903;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical sensory transduction protein ykog.
GN YKOG OR BSU13250.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RA Devine K.M.;
RT "Sequence of the Bacillus subtilis genome between xlyA and ykor.";
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
RW [2]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G., Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S., Borries R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S., Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M., Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A., Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T., Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D., Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi E., Galleron N., Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G., Guiseppe G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A., Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L., Joris B., Karamata D., Kaahara Y., Klaerr-Blanchard M., Klein C., Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M., Kurita K., Lepidou A., Lardinois S., Lauber J., Lazarevic V., Lee S.M., Levine A., Liu H., Masuda S., Maue C., Medigue C., Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M., Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H., Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M., Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S., Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadale Y., Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F., Sekiguchi J., Sakowska A., Seror S.J., Seror P., Shin B.S., Soldo B., Sorokin M., Tanakoshi A., Tanaka T., Takahashi H., Takemaru K., Takeuchi M., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A., Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A., Viari A., Wambut R., Wedler E., Wedler H., Weitzenecker T., Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K., Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;
RT "The complete genome sequence of the Gram-positive bacterium Bacillus subtilis";
RL Nature 390:249-256(1997).
CC -!- FUNCTION: Could be member of the two-component regulatory system ykoH/ykoG.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Potential).

ALIGNMENTS

34 395 35.2 227 1 PHOB_RHIME Q52990 rhizobium m
35 392 34.9 224 1 CIAR_STRPN Q54954 streptococc
36 392 34.9 231 1 PHOB_HAEIN P45189 haemophilus
37 391.5 34.9 220 1 VANR_ENTFA Q47744 enterococcu
38 390.5 34.8 226 1 YCBL_BACSU P42244 bacillus su
39 388 34.6 231 1 CTRA_CAUCR Q45994 caulobacter
40 385 34.3 232 1 CFXR_ECOLI P16244 escherichia
41 382 34.0 257 1 SPHR_SYNP7 P39663 synchococc
42 376.5 33.6 225 1 KDPE_ECOLI P21866 escherichia
43 373.5 33.3 229 1 PHOB_PSBAE P23620 pseudomonas
44 371 33.1 229 1 NISR_LACLA Q07597 lactococcus
45 368 32.8 217 1 CUTR_STRCO Q03756 streptomyce


```

CC -1- PTM: Phosphorylated by yKOH (Potential).
CC -1- SIMILARITY: Contains 1 response regulatory domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AJ002571; CAA05604.1; -.
CC EMBL; Z99110; CAB13182.1; -.
CC PIR; C69859; C69859.
CC HSP; P08402; I300.
CC Subtilisin; G013252; yKOH.
CC InterPro; IPR001789; Response reg.
CC InterPro; IPR001867; Trans reg C.
CC Pfam; PF00072; response_reg_1.
CC Pfam; PF00486; trans_reg_C; 1.
CC ProDom; PD000039; response_reg; 1.
CC ProDom; PD000329; Trans_reg_C; 1.
CC SMART; SM00448; REC; 1.
CC PROSITE; PS01110; RESPONSE REGULATORY; 1.
CC Hypothetical protein; Sensory transduction; Phosphorylation;
CC Transcription regulation; DNA-binding; Complete proteome.
CC DOMAIN 1 59 PHOSPHORYLATION (BY SIMILARITY).
FT MOD_RES 54 119
SQ SEQUENCE 228 AA; 26012 MW; F39A023E37ADBE32 CRC64;

Query Match 47.08; Score 527; DB 1; Length 228;
Best Local Similarity 48.48; Pred. No. 5.9e-33;
Matches 108; Conservative 50; Mismatches 57; Indels 8; Gaps 5;

Qy 4 ILIVDEQNARLFELELTHENYNVDTEYDGDGLKALSHYDYLDLILDLMLPSINGLEI 63
Db 6 ILIVDEKTRARVLQLELEYEGSVTIKINGTEGLDAAAEAGGYSLVLLDVMPLGSLGV 65

Qy 64 CRKIRQOOS-TPIIITAKSDTYDKVAGLDYGDADDYIVKFPDTEELLARTRAILRQPOK 122
Db 66 LRLRLKTDSTQPTVILLTARDSPDKVGLDIGANDYVTKPFEIEELLARTRAILRQNGTK 125

Qy 123 --DI--IDVNGITIDKNAPKVTNGAEIELTKTEYDLYLLAENKHNVMQREQLNHV 177
Db 126 TEDIGTFLTYDRLKVNKRETRVGRGKEVELTREFDLYVLMKHPQOVLTRQILSSW 185

Qy 178 GYNESEVTNVVDYVIRYLRNKLK-PYDRDKMIETVGVGVIR 219
Db 186 GFDYIGDNTNVVDYVIRYLRKLDYPEK-QLIHTRGVGYA 227

RESULT 2
RESID_BACSU STANDARD; PRT; 240 AA.
AC P35163;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DE 10-OCT-2003 (Rel. 42, Last annotation update)
GN RESD OR BSU23120.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168 / Marburg;
RA MEDLINE=95020538; PubMed=7934829;
RX Sorokin A.V., Zumstein E., Azevedo V., Ehrlich S.D., Serror P.;
RT "the organization of the Bacillus subtilis 168 chromosome region"
RL between the spoVA and serA genetic loci, based on sequence data.";
RN Mol. Microbiol. 10:385-395(1993).
RP [2]
RP SEQUENCE FROM N.A.

```

```

RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Betero M.G., Bessieres P., Bolotin A., Borchert S.,
RA Borriess R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Bruchsi C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA Denizot P., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,
RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Ghim S.Y., Glaser P., Goffeau A., Golligly E.J., Grandi G.,
RA Guisepi G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Maue C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetalle D., Porwollik S., Prescott A.M.,
RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivalta C., Roche E., Roche B., Rose M., Sadaie Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA Sekiguchi J., Sekowska A., Serror S.J., Serror P., Shin B.S., Soldo B.,
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tamakoshi A., Tanaka T., Terpetra P., Tognoni A.,
RA Viari A., Wambutt R., Vandenbol M., Vannier F., Vassarotti A.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.P., Zumstein E., Yoshikawa H., Danchin A.;
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
RT subtilis.";
RL Nature 390:249-256(1997).
RN [3]
RP FUNCTION.
RX MEDLINE=96200108; PubMed=8631715;
RA Sun G., Sharkova E., Chesnut R., Birkey S., Duggan M.F., Sorokin A.V.,
RA Pujic P., Ehrlich S.D., Hulett F.M.;
RT "Regulators of aerobic and anaerobic respiration in Bacillus
RT subtilis.";
RL J. Bacteriol. 178:1374-1385(1996).
RN [4]
RP FUNCTION.
RX MEDLINE=96272253; PubMed=8682783;
RA Nakano M.N., Zuber P., Glaser P., Danchin A., Hulett F.M.;
RT "Two-component regulatory proteins Resd-Rese are required for
RT transcriptional activation of fnr upon oxygen limitation in Bacillus
RT subtilis.";
RL J. Bacteriol. 178:3796-3802(1996).
CC -1- FUNCTION: MEMBER OF THE TWO-COMPONENT REGULATORY SYSTEM RESD/RESE.
CC REQUIRED FOR THE EXPRESSION OF RESA, CTAA, QCRABC AND FNR,
CC ACTIVATION ROLE IN GLOBAL REGULATION OF AEROBIC AND ANAEROBIC
CC RESPIRATION.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -1- PTM: PHOSPHORYLATED BY RESE (PROBABLE).
CC -1- SIMILARITY: Contains 1 response regulatory domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L09228; AAA67497.1; -.
CC EMBL; Z99116; CAB14244.1; -.
CC PIR; G69691; G69691.
CC HSP; P08402; I300.
CC Subtilisin; G013252; resd.
CC InterPro; IPR001789; Response reg.
CC InterPro; IPR001867; Trans reg C.
CC Pfam; PF00072; response_reg_1.

```



```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X83541; CAA58529.1; -.
DR PIR; S70164; S70164.
DR HSP; P08402; 1B00.
DR InterPro; IPR006291; Metal_resp_reg.
DR InterPro; IPR001785; response_reg.
DR InterPro; IPR001867; trans_reg_C.
DR Pfam; PF00072; response_reg; 1.
DR Pfam; PF00486; trans_reg_C; 1.
DR ProDom; PD000039; Response_reg; 1.
DR ProDom; PD000329; Trans_reg_C; 1.
DR SMART; SM00448; REC; 1.
DR TIGRFAMs; TIGR01387; cztR_silR_copR; 1.
DR PROSITE; PS0110; RESPONSE REGULATORY; 1.
KW Sensory transduction; Phosphorylation; Transcription regulation;
KW DNA-binding; Activator; Copper; Plasmid.
FT DOMAIN 1 117 RESPONSE REGULATORY.
FT MOD RES 52 52 PHOSPHORYLATION (BY SIMILARITY).
SQ SEQUENCE 226 AA; 25675 MW; 80502DAE2B23B23D CRC64;

Query Match 40.0%; Score 448.5; DB 1; Length 226;
Best Local Similarity 43.4%; Pred. No. 4.8e-27;
Matches 96; Conservative 39; Mismatches 83; Indels 3; Gaps 2;

Qy 1 MTQILIVEDQNARFLELTHENYNVDTEYDQGLDKALSHYDYDLIILDMPLPSING 60
Db 1 MQRILIVEDQKTRGLYQQGLVEEGYQADLFNNGRDGLGAASKGYDLIILDMPLDVG 60

Qy 61 LEICKIRQQ-QSTPIIITAKSDTYDKVAGLDYDGYVVKPDIIEELARIRAILRQ 119
Db 61 WQISALRESGHEBFLVFLAKONVRDKVGLGADYLIKPFDFELVARVTLRRA 120

Qy 120 PQK--DIIDVNGIITDKNAFKVTNGAEIELTKTEYDLYLLAENKNHVMQREQLNHV 177
Db 120 RSQAATVCTIADMTVDVRRVIRSGKKIHLTKGYVLELLQLTQCEVLPRLSISLVW 180

Qy 178 GYNEVETNVVDVYIRLNKLPDYDRDKMIETVRGVGYV 218
Db 181 NMNFDSTNVIDVAVRLRSKIDDDFPFKLIHTVRGAGYVL 221

RESULT 5
PRRA MYCTU STANDARD; PRT; 233 AA.
AC Q50136;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Transcriptional regulatory protein prrA.
GN PRRA OR ML2123 OR MLCB57.61C.
OS Mycobacterium leprae.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1769;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TN;
RX MEDLINE=21128732; PubMed=11234002;
RA Cole S.T., Biglener K., Parkhill J., James K.D., Thomson N.R.,
RA Wheeler P.R., Honore N., Garner T., Churcher C., Harris D.,
RA Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R.M., Devlin K., Duthoy S., Feltwell T., Fraser A., Hamlin N.,
RA Holroyd S., Hornsby T., Jagels K., Lacroix C., Maclean J., Moule S.,
RA Murphy L., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,
RA Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,
Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
Barrell B.G.;
RT "Massive gene decay in the leprosy bacillus.";
RL Nature 409:1007-1011(2001).
RN [2]
RP SEQUENCE OF 1-59 FROM N.A.
RA Smith D.R., Robison K.;
RA Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Member of the two-component regulatory system prrB/prrA.
CC Involved specifically in early intracellular multiplication of
CC Mycobacterium (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC -!- PTM: Phosphorylated by prrB (Potential).
CC -!- SIMILARITY: Contains 1 response regulatory domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z99494; CABI6701.1; -.
DR EMBL; AL583924; CAC31078.1; -.
DR EMBL; U15184; AAG3079.1; -.
DR PIR; T45359; T45359.
DR HSP; P08402; 1B00.
DR Leproma; ML2123; -.
DR InterPro; IPR001789; Response_reg.
DR InterPro; IPR001867; Trans_reg_C.
DR Pfam; PF00072; response_reg; 1.
DR Pfam; PF00486; trans_reg_C; 1.
DR ProDom; PD000039; Response_reg; 1.
DR ProDom; PD000329; Trans_reg_C; 1.
DR SMART; SM00448; REC; 1.
DR PROSITE; PS0110; RESPONSE REGULATORY; 1.
KW Sensory transduction; Phosphorylation; Transcription regulation;
KW DNA-binding; Complete proteome.
FT DOMAIN 9 123 RESPONSE REGULATORY.
FT MOD RES 58 58 PHOSPHORYLATION (BY SIMILARITY).
SQ SEQUENCE 233 AA; 25109 MW; A04D74E8537B6F1 CRC64;

Query Match 39.8%; Score 447; DB 1; Length 233;
Best Local Similarity 39.3%; Pred. No. 6.5e-27;
Matches 89; Conservative 52; Mismatches 76; Indels 6; Gaps 2;

Qy 3 QILIVDEQNARFLELTHENYNVDTEYDQGLDKALSHYDYDLIILDMPLPSINGLE 62
Db 9 RVLVDDSDVLASLGLSLGSGFVSTADGAELRNATETRPDAIVLDINPVLGVS 68

Qy 63 ICKIR-QQOSTPIIITAKSDTYDKVAGLDYDGYVVKPDIIEELARIRAILRQ-- 119
Db 69 VVTALRAMDNVDVCSVLSRSSVDDRVAGLEAGADYLVKPFVLAEIVARVKALLRGA 128

Qy 120 ---PQKDIIDVNGIITDKNAFKVTNGAEIELTKTEYDLYLLAENKNHVMQREQLNHV 176
Db 129 TATSSSETIATVGLPVDIPGRRARVNGVDVLTFRKFDLLVLAELHKTTLVLSRAQLDELV 180

Qy 177 WGYNEVETNVVDVYIRLNKLPDYDRDKMIETVRGVGYVIR 219
Db 189 WGYDPAADTNVDVDFIGYLRKLEANSFPRLHTVRGVGVLR 231

RESULT 6
PRRA MYCTU STANDARD; PRT; 233 AA.
AC Q10531;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Transcriptional regulatory protein prrA.
GN PRRA OR RV0903C OR MT0926 OR MTCY31.31C OR MB0927C.

```

OS Mycobacterium tuberculosis, and
OS Mycobacterium bovis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
NCBI_TaxID=1773, 1765;
[1]
RN
RP SEQUENCE FROM N.A.
RC SPECIES=M.tuberculosis; STRAIN=H37RV;
RX MEDLINE=9825987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeier K., Gao S., Barry C.E. III, Tekala F.,
RA Badcock K., Baaham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Felwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares R., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RL Nature 393:537-544 (1998).
[2]
RN
RP SEQUENCE FROM N.A.
RC SPECIES=M.tuberculosis; STRAIN=CDC 1551 / Oshkosh;
RX MEDLINE=22206494; PubMed=12218036;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J.F., DeBoy R., Dodson R., Gwinn M., Haft D., Hickey E.,
RA Kolonay J.J., Nelson W.C., Onyon L.A., Ermolaeva M., Salzberg S.L.,
RA Delcher A., Uitterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W., Jacobs W.R. Jr., Venter J.C., Fraser C.M.;
RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";
RL J. Bacteriol. 184:5479-5490(2002).
[3]
RN
RP SEQUENCE FROM N.A.
RC SPECIES=M.bovis; STRAIN=AF2122/97;
RX MEDLINE=22709107; PubMed=12788972;
RA Garnier T., Eiglmeier K., Camus J.-C., Medina N., Mansoor H.,
RA Pryor M., Duthoy S., Grondin S., Lacroix C., Monsemp C., Simon S.,
RA Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
RA Parkhill J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
RT "The complete genome sequence of Mycobacterium bovis.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).
[4]
RN
RP FUNCTION.
RC SPECIES=M.tuberculosis; STRAIN=MC103;
RX MEDLINE=21950537; PubMed=11953357;
RA Ewann F., Jackson M., Pethe K., Cooper A., Mielcarek N.,
RA Ensergueux D., Gicquel B., Lochet C., Supply P.;
RT "Transient requirement of the PrfA-PrfB two-component system for
RT early intracellular multiplication of Mycobacterium tuberculosis.";
RL Infect. Immun. 70:2256-2263(2002).
CC -!- FUNCTION: Member of the two-component regulatory system prfB/prfA.
CC Involved specifically in early intracellular multiplication of
CC Mycobacterium.
CC
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC -!- PTM: Phosphorylated by prfB (Potential).
CC -!- SIMILARITY: Contains 1 response regulatory domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

DR EMBL: Z73101; AAA97378.1; ALT INIT.
DR EMBL: AB006979; AAK45173.1; ALT INIT.
DR EMBL: BX248337; CAD93788.1; ALT_INIT.
DR PIR: D70783; D70783.
DR HSP: P08402; I800.
DR TIGR: MT0926; -.
DR Tuberculin; RV0903c; -.
DR

RA Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
RA Guiseppe G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hooson S., Kluett M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klier-Blanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Maue C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA Neone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivalta C., Roche E., Roche B., Rose M., Sadaie Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,
RA Srokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tanakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zmestein E., Yoshikawa H., Danchin A.,
RT "The complete genome sequence of the Gram-positive bacterium *Bacillus*
RT *subtilis*.";
RL Nature 390:249-256(1997).
RN [4]
RN FUNCTION, AND VARIANT JH17041.
RC STRAIN=168 / JH642;
RX MEDLINE=99047586; PubMed=9829949;
RA Fabret C., Hoch J.A.;
RT "A two-component signal transduction system essential for growth of
RT *Bacillus subtilis*: implications for anti-infective therapy.";
RL J. Bacteriol. 180:6375-6383(1998).
RN [5]
RN FUNCTION, AND MUTAGENESIS OF ASP-53.
RP STRAIN=168;
RX MEDLINE=20340961; PubMed=10878122;
RA Fukuchi K., Kasahara Y., Asai K., Kobayashi K., Moriya S.,
RA Ogasawara N.;
RT "The essential two-component regulatory system encoded by *yycF* and
RT *yycG* modulates expression of the *ftsAZ* operon in *Bacillus subtilis*.";
RL Microbiology 146:1573-1583(2000).
CC -1- FUNCTION: Member of the two-component regulatory system *yycG/yycF*
CC involved in the regulation of the *ftsAZ* operon. Binds to the *ftsAZ*
CC P1 promoter sequence in vitro.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -1- DEVELOPMENTAL STAGE: Expressed during exponential growth and shut
CC down at the entry into stationary phase.
CC -1- PTM: Phosphorylated by *yycG* (Probable).
CC -1- SIMILARITY: Contains 1 response regulatory domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D26185; BAA05173.1; -;
DR EMBL; D78193; BAA1300.1; -;
DR EMBL; Z99124; CAB16078.1; -;
DR PIR; S65967; S65967.
DR HSP; P08402; 1800.
DR Subtilist; BG10001; *yycF*.
DR InterPro; IPR001789; Response_reg.
DR InterPro; IPR001867; Trans_reg_C.
DR Pfam; PF00072; response_reg; 1.
DR Pfam; PF00486; trans_reg_C; 1.
DR ProDom; PD000039; Response_reg; 1.
DR ProDom; PD000329; Trans_reg_C; 1.
DR SMART; SM00448; REC; 1.
DR PROSITE; PS0110; RESPONSE REGULATORY; 1.
DR Sensory transduction; Phosphorylation; Transcription regulation;
KW DNA-binding; Complete proteome.

FT DOMAIN 1 117 RESPONSE REGULATORY.
FT MOD_RES 53 53 PHOSPHORYLATION (PROBABLE).
FT VARIANT 215 215 H -> P (IN JH17041; THERMOSENSITIVE).
FT MUTAGEN 53 53 D -> H: CONSTITUTIVELY ACTIVE.
SQ SEQUENCE 235 AA; 27226 MW; 201BC21CBB5EC6B4 CRC64;
Query Match 39.7%; Score 445; DB 1; Length 235;
Best Local Similarity 41.4%; Pred. No. 9.3e-27;
Matches 94; Conservative 43; Mismatches 80; Indels 10; Gaps 2;
QY 3 QILVEDEQNARFLLELTHENYNVDYDGDGDLKALSHYDYLIIIDLMPLSNGLE 62
DB 4 KILVDDKPIADILFNLKKEGYVHCADNGEAVEMVEELQPDILIDMLPNKDGVE 63
QY 63 ICRKIPQQQSTPIIIITAKSDTYDKVAGLDYGDADDYIVKPFDEIELLARIRALRQ--- 119
DB 64 VCREVKKYDMPIMLTAKDSEIDKVIKLEIGADDYVTRPFSTRELLARVKNLRQLTT 123
QY 120 -----PQKDIIDVNGITIDKNAFKVTVNGAEITKTYDYLLYLLAENKNHVMQREQIL 173
DB 124 APAEPEPSNEHIGSLVFPDAYVVKRDEITELTHREFFELLYLAKHIGVMTREHLL 183
QY 174 NHVWGYNSEVTNVVDYIRYLNLKPKY-DRDKMIETVRGVGVYIR 219
DB 184 QTWVGVDYFGDVRTVDVTVRLREKIEDNPNHWNIVTRRGVGYLR 230
RESULT 8
MTR_A MYCPA STANDARD; PRT; 225 AA.
AC Q93CB8;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE DNA-binding response regulator mtrA.
GN MTR_A.
OS Mycobacterium paratuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1770;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 19698;
RA Urbanic K.W., Mutharia L.M.;
RT "Identification and initial characterization of the mtrAB two-
RT component signal transduction system of *Mycobacterium avium* subspecies
RT paratuberculosis.";
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Member of the two-component regulatory system mtrA/mtrB
CC (by similarity).
CC -1- PTM: Phosphorylated by mtrB (Probable).
CC -1- SIMILARITY: Contains 1 response regulatory domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF410884; AAL10207.1; -;
DR InterPro; IPR001789; Response_reg.
DR InterPro; IPR001867; Trans_reg_C.
DR Pfam; PF00072; response_reg; 1.
DR Pfam; PF00486; trans_reg_C; 1.
DR ProDom; PD000039; Response_reg; 1.
DR ProDom; PD000329; Trans_reg_C; 1.
DR SMART; SM00448; REC; 1.
DR PROSITE; PS0110; RESPONSE REGULATORY; 1.
DR Sensory transduction; Phosphorylation; Transcription regulation;
KW DNA-binding.
FT DOMAIN 1 117 RESPONSE REGULATORY.

FT MOD_RES 53 53 PHOSPHORYLATION (BY SIMILARITY).
 SQ SEQUENCE 225 AA; 24922 MW; A195E62B3D42DEE3 CRC64;
 Query Match 39.6%; Score 444.5; DB 1; Length 225;
 Best Local Similarity 42.4%; Pred. No. 9.6e-27;
 Matches 92; Conservative 44; Mismatches 78; Indels 3; Gaps 2;
 QY 3 QILVDEONLARFLELELTHENVYDYGDDGLDKALSHYDYLILDLMLPSINGLE 62
 Db 4 RILVDDASLAEMLTIVLRGEGFTAVIGDGTQALTAVRELRLPDLVLLDLMLPGMGID 63
 QY 63 ICRKIROQOSTPIIIITAKSDTYDKVAGLDYGADDYIVKPPDIEELLARIRAILRRQPK 122
 Db 64 VCRVLRADSGVPIVWLTKTDVTVLVGLSGADDYIMKPKPKELVARVRLRRNDE 123
 QY 123 --DIIDVNGITIDKNAFKVTNGAEIETKTGYDLYLLAENKHNVMOREQIILNHVWYN 180
 Db 124 PAEMLSIADVEIDVPAHKVTRNGEQISLTPLEFDLLVALARKPRQVTRDVLLEQVWGYR 183
 QY 181 SEVTNVVDVYIRYLRNKL-KPYDRDKMIETVRGVGY 216
 Db 184 HPADTRLNVHVQRLRAKVEKDPENPTVTLTVRGVGY 220

RESULT 9

MTRA_MYCTU
 ID MYCTU STANDARD; PRT; 225 AA.
 AC Q50447;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE DNA-binding response regulator mtrA.
 GN MTRA OR RV3246C OR MT3344 OR MTCY20B11.21C OR MB3274C.
 OS Mycobacterium tuberculosis, and
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1773, 1765;
 RX [1]
 RP SEQUENCE FROM N.A. AND CHARACTERIZATION.
 RC SPECIES=M.tuberculosis; STRAIN=H37RV;
 RX MEDLINE=98295987; PubMed=8655513;
 RA Via L.E., Curcio R., Mudd M.H., Dhandayuthapani S., Ulmer R.J.,
 RA Deretic V.;
 RA "Elements of signal transduction in Mycobacterium tuberculosis: in
 RT vitro phosphorylation and in vivo expression of the response regulator
 RT MtrA."
 RL J. Bacteriol. 178:3314-3321(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC SPECIES=M.tuberculosis; STRAIN=H37RV;
 RX MEDLINE=98295987; PubMed=9634230;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 RA Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies K., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
 RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
 RA Oliver S., Oeborne J., Quail M.A., Rajandream M.A., Rogers J.,
 RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
 RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
 RT "Deciphering the biology of Mycobacterium tuberculosis from the
 RT complete genome sequence."
 RL Nature 393:537-544(1998).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC SPECIES=M.tuberculosis; STRAIN=CDC 1551 / Oshkosh;
 RX MEDLINE=2206494; PubMed=12218036;
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 RA Peterson J., DeBoy R., Dodson R., Gwinn M., Haft D., Hickey E.,
 RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M., Salzberg S.L.,
 RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
 RA Bishai W., Jacobs W.R. Jr., Venter J.C., Fraser C.M.;
 RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and

laboratory strains."
 RL J. Bacteriol. 184:5479-5490(2002).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC SPECIES=M.bovis; STRAIN=AF2122/97;
 RX MEDLINE=22709107; PubMed=12788972;
 RA Garnier T., Eigmeier K., Camus J.-C., Medina N., Mansoor H.,
 RA Pryor M., Duthoy S., Grondin S., Lacroix C., Monsemp C., Simon S.,
 RA Harris B., Akin R., Doggett J., Maves R., Keating L., Wheeler P.R.,
 RA Parkhill J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
 RT "The complete genome sequence of Mycobacterium bovis."
 RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).
 CC -!- FUNCTION: Member of the two-component regulatory system mtrA/mtrB.
 CC -!- PTM: Phosphorylated by mtrB (Probable).
 CC -!- SIMILARITY: Contains 1 response regulatory domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.ebi.ac.uk/announcements>)
 CC or send an email to license@isb-sib.ch.
 CC -----
 DR EMBL; U01971; AAB07804.1; -;
 DR EMBL; Z95121; CAB08347.1; ALT INIT.
 DR EMBL; AE007145; AAK47686.1; -;
 DR EMBL; BX248345; CAD95366.1; ALT_INIT.
 DR PIR; H70592; H70592.
 DR HSSP; P08402; I800.
 DR TIGR; MT3344; -;
 DR TubercuList; RV3246C; -;
 DR InterPro; IPR001789; Response reg.
 DR InterPro; IPR001867; Trans_reg_C.
 DR Pfam; PF00072; response_reg; 1.
 DR Pfam; PF00486; trans_reg_C; 1.
 DR ProDom; PD000039; Response reg; 1.
 DR ProDom; PD000329; Trans_reg_C; 1.
 DR SMART; SM00448; REC; 1.
 DR PROSITE; PS0110; RESPONSE_REGULATORY; 1.
 DR Sensory transduction; Phosphorylation; Transcription regulation;
 KW DNA-binding; Complete proteome.
 FT DOMAIN 1 117 RESPONSE_REGULATORY (BY SIMILARITY).
 FT MOD_RES 53 53 PHOSPHORYLATION (BY SIMILARITY).
 SQ SEQUENCE 225 AA; 24932 MW; A345B562B3D42DEE3 CRC64;

Query Match 39.6%; Score 444.5; DB 1; Length 225;
 Best Local Similarity 42.4%; Pred. No. 9.6e-27;
 Matches 92; Conservative 44; Mismatches 78; Indels 3; Gaps 2;

QY 3 QILVDEONLARFLELELTHENVYDYGDDGLDKALSHYDYLILDLMLPSINGLE 62
 Db 4 RILVDDASLAEMLTIVLRGEGFTAVIGDGTQALTAVRELRLPDLVLLDLMLPGMGID 63
 QY 63 ICRKIROQOSTPIIIITAKSDTYDKVAGLDYGADDYIVKPPDIEELLARIRAILRRQPK 122
 Db 64 VCRVLRADSGVPIVWLTKTDVTVLVGLSGADDYIMKPKPKELVARVRLRRNDE 123
 QY 123 --DIIDVNGITIDKNAFKVTNGAEIETKTGYDLYLLAENKHNVMOREQIILNHVWYN 180
 Db 124 PAEMLSIADVEIDVPAHKVTRNGEQISLTPLEFDLLVALARKPRQVTRDVLLEQVWGYR 183
 QY 181 SEVTNVVDVYIRYLRNKL-KPYDRDKMIETVRGVGY 216
 Db 184 HPADTRLNVHVQRLRAKVEKDPENPTVTLTVRGVGY 220

RESULT 10

MTRA_MYCTE
 ID MYCTE STANDARD; PRT; 225 AA.
 AC Q9CCJ2;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)


```
CC -----
DR EMBL; AF005358; AAB92482.1; -.
DR HSP; P08402; I800.
DR InterPro; IPRO06291; Metal_resp_reg.
DR InterPro; IPRO01789; Response_reg.
DR InterPro; IPRO01867; Trans_reg_C.
DR Pfam; PF00072; response_reg_1.
DR Pfam; PF00486; trans_reg_C; 1.
DR ProDom; PD000039; Response_reg; 1.
DR ProDom; PD000329; Trans_reg_C; 1.
DR SMART; SMO0448; REC; 1.
DR TIGRFAMs; TTGR01387; cztr_silr_copr; 1.
DR PROSITE; PS0110; RESPONSE_REGULATORY; 1.
KW Sensory transduction; Phosphorylation; Transcription regulation;
KW DNA-binding; Activator; Zinc; Cadmium.
FT DOMAIN 1 115 PHOSPHORYLATION (BY SIMILARITY).
FT MOD RES 51 51 RESPONSE REGULATION.
SQ SEQUENCE 229 AA, 25971 MW, 9D89D1BF6E62E98C CRC64;

Query Match      38.9%; Score 437; DB 1; Length 229;
Best Local Similarity 41.7%; Pred.No. 3.6e-26;
Matches 91; Conservative 42; Mismatches 83; Indels 2; Gaps 3;

QY 3 QILLVEDEQNLAFLLELTTHENVVDTEYDGODGLDKALSHYYDLIILDMLPSINGLE 62
   :|||||::||::||::||::||::||::||::||::||::||::||::||::||:
DB 2 RLIIIVEDEPKTGMVKRKLTEAGFIADWVEDGVTVGLHQAETEEYDYLIIDVLPFGHDGWT 61

QY 63 ICCKIRQQOQTPIIIITAKSDTYKVGAGLDYGADDIVTKPFDIEELIARTAILRRQPQK 122
   :::::::||::||::||::||::||::||::||::||::||::||::||::||:
DB 62 VLRELRAHSPTVLFIARDDDGVRKVGLGADDDVVKKFPDFVELVARVSILRRQAR 121

QY 123 D--IIDVNGITIDKNAPKVTNVAEGIELTKYTEYDLLLYLLAENKHNVQREQLINHWGYN 180
DB 122 ESTVLRIADLEDLTRRKATRGDVLLTAKEFALLWLMMRREGIELPRATIASQVWDWN 181

QY 181 SEVETNWVDVIYRLNKLKPYDRDKMIETVRGVGI 218
   :||||::||::||::||::||::||::||::||::||:
DB 182 FNSDTNVNDAAIRLRASKIDDAYEPKLIHTVRGMGYVL 219

RESULT 13
YEDW_ECOLI
ID YEDW_ECOLI STANDARD; PRT; 223 AA.
AC P76340; P97172;
DT 01-NOV-1997 (Rel. 35, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Probable transcriptional regulatory protein yedw.
GN YEDW OR B1969.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OX Enterobacteriaceae; Escherichia.
ON NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RA "The complete genome sequence of Escherichia coli K-12." ;
RL Science 277:1453-1474 (1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=97251358; PubMed=9097040;
RA Itoh T., Aiba H., Baba T., Fujita K., Hayashi K., Inada T.,
RA Isano K., Kasai H., Kimura S., Kitakawa M., Kitagawa M.,
RA Makino S., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K.,
RA Nakade S., Nakamura Y., Nashimoto H., Nishio Y., Oshima T.,
RA Saito N., Sampei G., Seki Y., Siwasundaram S., Tagami H.,
RA Takeda J., Takemoto K., Wada C., Yamamoto Y., Horiiuchi T.;
```

```

RT "A 460-kb DNA sequence of the Escherichia coli K-12 genome
RL corresponding to the 40.1-50.0 min region on the linkage map.";
CC DNA Res. 3:379-392(1996).
CC -1- FUNCTION: PROBABLE MEMBER OF A TWO-COMPONENT REGULATORY SYSTEM
CC YEDV/YEDV.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -1- PTM: PHOSPHORYLATED BY YEDV (POTENTIAL).
CC -1- SIMILARITY: Contains 1 response regulatory domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC ENBL; AE000288; AAC75035.1; ALT INIT.
DR ENBL; D90835; BAA15796.1; ALT INIT.
DR ENBL; D90836; BAA15797.1; ALT_INIT.
DR HSP; P08402; I300.
DR EcoGene; EG14045; yedV.
DR InterPro; IPR006291; Metal resp reg.
DR InterPro; IPR001789; Response reg.
DR InterPro; IPR001867; Trans reg_C.
DR Pfam; PF00072; response_reg_1.
DR Pfam; PF00486; trans_reg_C_1.
DR ProDom; PD000039; Response_reg_1.
DR ProDom; PD000329; Trans_reg_C_1.
DR SMART; SM00448; REC; 1.
DR TIGRfams; TIGR01387; cztr slir copr; 1.
DR TRUST; PS01110; RESPONSE REGULATORY; 1.
KW Hypothetical protein; Sensory transduction; Phosphorylation;
KW Transcription regulation; DNA-binding; Complete proteome.
FT DOMAIN 1 115 RESPONSE REGULATORY.
FT MOD RES 51 51 PHOSPHORYLATION (BY SIMILARITY).
SQ SEQUENCE 223 AA; 25017 MW; 4A595D7430C2E2B5 CRC64;

Query Match 38.4%; Score 430.5; DB 1; Length 223;
Best Local Similarity 40.9%; Pred. No. 1.le-25;
Matches 88; Conservative 46; Mismatches 80; Indels 1; Gaps 1;

Qy 3 QILIVDEONLARFLELELTHENNVVDTEVDGQGLDKALSHYVLIILDLPLPSINGLE 62
Db 2 KILLIENQRTQBTWGTGLSEAGVVIDAVSDGRDGLYALKDDYALILDLIMLPMDGMQ 61

Qy 63 ICRKIROQQSTPIIIITAKSDTYDKVAGLDYGDADDYIVKPFIDIEELIARILRR-QPQ 121
Db 62 ILQTLRTAKQTPVICTARSDVDRVRLSGANDYLVKPFSELLARVRAQLRQHALL 121

Qy 122 KDIIDVNGITIDKNAFKVTNGAEIETKTEYDLYLLAENKHNVMQREQLNHNWGYNS 181
Db 122 NSTLEISGLRMSVSHSVSRDINSITLTKREFQLLMLASRAGEIIPRTVIASEINGINF 181

Qy 182 EVENVVDVIVIRYLNKLKPYDRDKMLETVRGVGY 216
Db 182 DSDTNTVDVIVIRLURAKVDDPFPEKLITATIRGMGY 216

RESULT 14
PHOP_BACSU STANDARD; PRT; 240 AA.
AC P13792; O34804;
DT 01-JAN-1990 (Rel. 13, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Alkaline phosphatase synthesis transcriptional regulatory protein
DE PHOP.
GN PHOP OR BSU29110.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN=168 / JH642;
RX MEDLINE=93087210; PubMed=1454550;
RA Lee J.W., Hulett F.M.;
RT "Nucleotide sequence of the phoP gene encoding PhoP, the response
RL regulator of the phosphate regulon of Bacillus subtilis.";
RL Nucleic Acids Res. 20:5848-5848(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=87250247; PubMed=3036763;
RA Seki T., Yoshikawa H., Takahashi H., Saito H.;
RT "Cloning and nucleotide sequence of phoP, the regulatory gene for
RT alkaline phosphatase and phosphodiesterase in Bacillus subtilis.";
RL J. Bacteriol. 169:2913-2916(1987).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98048467; PubMed=9387221;
RA Lapidus A., Galleron N., Sorokin A., Ehrlich S.D.;
RT "Sequencing and functional annotation of the Bacillus subtilis genes
RT in the 200 kb rrmB-dnaB region.";
RL Microbiology 143:3431-3441(1997).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA Borriess R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,
RA Entian K.D., Errington J., Fabret C., Ferrati E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Ghim S.Y., Glaser P., Goffeau A., Golightly B.J., Grandi G.,
RA Guisepi G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Maue C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Porvollik S., Prescott A.M.,
RA Prescan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadale Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA Sekiguchi J., Sekowska A., Seror S.J., Seror P., Shin B.S., Soldo B.,
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tanakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.P., Zumbstein E., Yoshikawa H., Danchin A.,
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
RL subtilis.";
RL Nature 390:249-256(1997).
RN [5]
RP SEQUENCE OF 212-240 FROM N.A.
RX MEDLINE=89053932; PubMed=3142862;
RA Seki T., Yoshikawa H., Takahashi H., Saito H.;
RT "Nucleotide sequence of the Bacillus subtilis phoR gene.";
RL J. Bacteriol. 170:5935-5938(1988).
CC -1- FUNCTION: MEMBER OF THE TWO-COMPONENT REGULATORY SYSTEM PHOP/PHOR
CC INVOLVED IN THE REGULATION OF ALKALINE PHOSPHATASE GENES PHOA AND
CC PHOB AND OF PHOSPHODIESTERASE.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -1- PTM: Phosphorylated by phoR (Probable).
CC -1- SIMILARITY: Contains 1 response regulatory domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 4, 2004, 17:59:28 ; Search time 48.3761 Seconds
(without alignments)
1428.359 Million cell updates/sec

Title: US-09-006-627-4

Perfect score: 1122

Sequence: 1 MTQLIVDEQNLFARLE.....KPYDRDKMETVRGVGYVIR 219

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_25.*

1: sp_archaea.*

2: sp_bacteria.*

3: sp_fungi.*

4: sp_human.*

5: sp_invertebrate.*

6: sp_mammal.*

7: sp_mhc.*

8: sp_organelle.*

9: sp_phase.*

10: sp_plant.*

11: sp_rodent.*

12: sp_virus.*

13: sp_vertebrate.*

14: sp_unclassified.*

15: sp_rvirus.*

16: sp_bacteriap.*

17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1122	100.0	219	Q9KJN4	Q9KJN4 staphylococ
2	968	86.3	219	Q8CP82	Q8CP82 staphylococ
3	656	58.5	133	Q99U72	Q99U72 staphylococ
4	649	57.8	228	Q836Q8	Q836Q8 enterococc
5	646	57.6	228	Q88WS4	Q88WS4 lactobacill
6	628	56.0	226	Q92BX8	Q92BX8 listeria mo
7	588.5	52.5	229	Q8RB10	Q8RB10 thermoanaer
8	583	52.0	230	Q9CF87	Q9CF87 lactococcus
9	581	51.8	230	Q9K570	Q9K570 lactococcus
10	579	51.6	226	Q9RP20	Q9RP20 listeria mo
11	574.5	51.2	229	Q8E3T4	Q8E3T4 streptococc
12	574.5	51.2	229	Q8DY68	Q8DY68 streptococc
13	574	51.2	228	Q87527	Q87527 streptococc
14	570.5	50.8	230	Q9S151	Q9S151 streptococc
15	568	50.6	228	Q8P2J8	Q8P2J8 streptococc
16	566	50.4	228	Q878B2	Q878B2 streptococc

17	560	49.9	228	2	Q9RQL2	Q9RQL2 streptococc
18	552.5	49.2	229	16	Q8ESN2	Q8ESN2 oceanobacil
19	539	48.0	222	16	Q9KFU9	Q9KFU9 bacillus ha
20	526.5	46.9	228	16	Q8YXL8	Q8YXL8 anabaena sp
21	504.5	45.0	233	16	Q8RE63	Q8RE63 fusobacteri
22	501.5	44.7	240	16	Q8R9H7	Q8R9H7 thermoanaer
23	496.5	44.3	246	16	Q89917	Q89917 clostridium
24	494	44.0	227	16	Q81R46	Q81R46 bacillus an
25	490.5	43.7	232	16	Q971B8	Q971B8 clostridium
26	489.5	43.6	228	16	Q8XHJ5	Q8XHJ5 clostridium
27	488.5	43.5	225	16	Q9RWC6	Q9RWC6 deinococcus
28	487	43.4	242	16	Q88YL3	Q88YL3 lactobacill
29	486.5	43.4	224	16	Q81NG7	Q81NG7 bacillus an
30	482.5	43.0	224	16	Q55733	Q55733 synecocyst
31	481.5	42.9	224	16	Q81BG1	Q81BG1 bacillus ce
32	481.5	42.9	231	16	Q8XMP6	Q8XMP6 clostridium
33	481	42.9	238	16	Q9KCU1	Q9KCU1 bacillus ha
34	479	42.7	238	16	Q92A60	Q92A60 listeria in
35	478.5	42.6	231	16	Q8EPB3	Q8EPB3 oceanobacil
36	476	42.4	238	16	Q8Y5V7	Q8Y5V7 listeria mo
37	475.5	42.4	223	16	Q81YR2	Q81YR2 bacillus an
38	475	42.3	239	16	Q9K850	Q9K850 bacillus ha
39	472	42.1	243	16	Q82S51	Q82S51 streptomyce
40	468.5	41.8	232	16	Q81147	Q81147 bacillus ce
41	468	41.7	224	16	Q8XS15	Q8XS15 ralstonia s
42	468	41.7	238	16	Q81S26	Q81S26 bacillus an
43	467	41.6	238	16	Q81FU1	Q81FU1 bacillus ce
44	466	41.5	234	16	Q882M8	Q882M8 lactobacill
45	465.5	41.5	232	16	Q895Y5	Q895Y5 clostridium

ALIGNMENTS

RESULT 1

Q9KJN4	PRELIMINARY;	PRT;	219 AA.
ID	Q9KJN4		
AC	Q9KJN4;		
DC	01-OCT-2000 (TremBLrel. 15, Created)		
DT	01-OCT-2000 (TremBLrel. 15, Last sequence update)		
DE	Putative response regulator ArlR (Truncated).		
OS	ARLR OR TRUNCATED-ARLR OR SAVI415 OR MWI305.		
GN	Staphylococcus aureus (strain Mu50 / ATCC 700699),		
OS	Staphylococcus aureus, and		
OC	Staphylococcus aureus (strain MW2).		
OC	Bacteria; Firmicutes; Bacillales; Staphylococcus.		
OX	NCBI_TaxID=158878, 1280, 196620;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=ISP794;		
RX	MEDLINE=20327575; PubMed=10869073;		
RA	Fournier B., Hooper D.C.;		
RT	"A new two-component regulatory system involved in adhesion,		
RT	autolysis, and extracellular proteolytic activity of Staphylococcus		
RT	aureus.";		
RL	J. Bacteriol. 182:3955-3964 (2000).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=Mu50;		
RX	MEDLINE=21311952; PubMed=11418146;		
RA	Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,		
RA	Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,		
RA	Kanamori M., Macemaru H., Maruyama A., Murakami H., Hoshoya A.,		
RA	Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,		
RA	Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,		
RA	Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,		
RA	Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;		
RT	"Whole genome sequencing of methicillin-resistant Staphylococcus		
RT	aureus.";		
RL	Lancet 357:1225-1240 (2001).		
RN	[3]		
RP	SEQUENCE FROM N.A.		

DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0000156; F:two-component response regulator activity; IEA.
DR GO; GO:0007600; P:sensory perception; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR GO; GO:0000160; P:two-component signal transduction system (p. . .); IEA.
DR InterPro; IPR001789; Response reg.
DR InterPro; IPR005829; Sug. transporter.
DR Pfam; PF00072; response_reg; 1.
DR ProDom; PD000039; Response_reg; 1.
DR SMART; SM00448; REC; 1.
DR PROSITE; PS0110; RESPONSE_REGULATORY; 1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW Phosphorylation; Sensory transduction; Complete proteome.
SQ SEQUENCE 133 AA; 15309 MW; 207099ABADA2DB3F CRC64;

Query Match 58.5%; Score 656; DB 16; Length 133;
Best Local Similarity 100.0%; Pred. No. 2.3e-42;
Matches 129; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTQILIVDEQNARFLELELTHENVNVDTEYDQGLDKALSHYDYLIILDLMLPSING 60
DB 1 MTQILIVDEQNARFLELELTHENVNVDTEYDQGLDKALSHYDYLIILDLMLPSING 60

QY 61 LEICRKIRQQQSTPIIIITAKSDTYDKVAGLDYGADDYIVKFPDIEELLARIRAILRRQP 120
DB 61 LEICRKIRQQQSTPIIIITAKSDTYDKVAGLDYGADDYIVKFPDIEELLARIRAILRRQP 120

QY 121 OKDIIDVNG 129
DB 121 OKDIIDVNG 129

RESULT 4
Q836Q8 ID Q836Q8 PRELIMINARY; PRT; 228 AA.
AC Q836Q8;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE DNA-binding response regulator.
GN EF1050.
OS Enterococcus faecalis (Streptococcus faecalis).
OC Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.
OX NCBI_TaxID=1351;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=V583 / ATCC 700802;
RX MEDLINE=22550857; PubMed=12663927;
RA Paulsen I.T., Banerjee L., Myers G.S.A., Nelson K.E., Seshadri R.,
RA Read T.D., Fouts D.E., Eisen J.A., Gill S.R., Heidelberg J.F.,
RA Tettelin H., Dodson R.J., Umayam L., Brinkac L., Beanan M.,
RA Daugherty S., DeBoy R.T., Durkin S., Kolonay J., Madupu R., Nelson W.,
RA Vamathevan J., Tran B., Upton J., Hansen T., Shetty J., Khouri H.,
RA Utterback T., Radune D., Ketchum K.A., Dougherty B.A., Fraser C.M.,
RT "Role of mobile DNA in the evolution of vancomycin-resistant
RT Enterococcus faecalis".
RL Science 299:2071-2074(2003).
DR EMBL; AE016950; AAC00853.1; --
DR TIGR; EF1050; --
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0000156; F:two-component response regulator activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR GO; GO:0000160; P:two-component signal transduction system (p. . .); IEA.
DR InterPro; IPR001789; Response reg.
DR InterPro; IPR005829; Sug. transporter.
DR Pfam; PF00072; response_reg; 1.
DR Pfam; PF00486; trans_reg_C; 1.
DR ProDom; PD000039; Response_reg; 1.

DR ProDom; PD000039; Trans_reg_C; 1.
DR SMART; SM00448; REC; 1.
DR PROSITE; PS0110; RESPONSE_REGULATORY; 1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW DNA-binding; Complete proteome.
SQ SEQUENCE 228 AA; 26455 MW; 6D5E2F6C8B14CB8C CRC64;

Query Match 57.8%; Score 649; DB 16; Length 228;
Best Local Similarity 55.4%; Pred. No. 1.5e-41;
Matches 129; Conservative 38; Mismatches 46; Indels 20; Gaps 2;

QY 1 MTQILIVDEQNARFLELELTHENVNVDTEYDQGLDKALSHYDYLIILDLMLPSING 60
DB 1 MSNLIIEDEKNARFVELELKHEGYTEHYNGRTGLEAALNNEWDAILDLMLPELNG 60

QY 61 LEICRKIRQQQSTPIIIITAKSDTYDKVAGLDYGADDYIVKFPDIEELLARIRAILRRQP 120
DB 61 LEICRKIRQQQSTPIIIITAKSDTYDKVAGLDYGADDYIVKFPDIEELLARIRAILRRQP 120

QY 121 OKDIIDVNG 129
DB 121 OKDIIDVNG 129

QY 167 MOREQILNHWGYNSEVETNVDYIRYLKLPYDRDKMIETVGVGYVIR 219
DB 175 LARDVLLNKNWGYETEVETNVDYIRYLKRNKIDVPGEESYIQTVRGTGYMR 227

RESULT 5
Q88WS4 ID Q88WS4 PRELIMINARY; PRT; 228 AA.
AC Q88WS4;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Response regulator.
GN RRP5 OR LP 1544.
OC Lactobacillus plantarum.
OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
OC Lactobacillus.
OX NCBI_TaxID=1590;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCIMB 8826 / WCFS1;
RX MEDLINE=22480296; PubMed=12566566;
RA Kleerebezem M., Boekhorst J., van Kranenburg R., Molenaar D.,
RA Kuipers O.P., Leer R., Tarchini R., Peters S.A., Sandbrink H.M.,
RA Fiers M.W.E.J., Stiekema W., Klein Lankhorst R.M., Bron P.A.,
RA Hoffer S.M., Nierop Groot M.N., Kerkhoven R., De Vries M., Ursing B.,
RA De Vos W.M., Siezen R.J.;
RT "Complete genome sequence of Lactobacillus plantarum WCFS1";
RL Proc. Natl. Acad. Sci. U.S.A. 100:1990-1995(2003).
DR EMBL; AL935256; CA063994.1; --
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0000156; F:two-component response regulator activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR GO; GO:0000160; P:two-component signal transduction system (p. . .); IEA.
DR InterPro; IPR001789; Response reg.
DR InterPro; IPR005829; Sug. transporter.
DR InterPro; IPR001867; Trans reg C.
DR Pfam; PF00072; response_reg; 1.
DR Pfam; PF00486; trans_reg_C; 1.
DR PROSITE; PS0110; RESPONSE_REGULATORY; 1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW Complete proteome.
SQ SEQUENCE 228 AA; 26372 MW; 54C48DE8A4DE488A CRC64;

Query Match 57.6%; Score 646; DB 16; Length 228;
Best Local Similarity 55.9%; Pred. No. 2.6e-41;
Matches 127; Conservative 38; Mismatches 54; Indels 8; Gaps 1;


```

QY 1 MTQILIVEDQNLFARFLELELTHENVNVDTEYDQDGLDKALSHYVDLIILDLMLPSING 60
Db 1 MSRLIIVEDKNLFARFVELELKHGVDIQVEYNGRKGDLDALEDFDAILDLMLPELNG 60

QY 61 LEICRKIROOQSTPIIIITAKSDTYDKVAGLDYGADDDYIVKFPDIBELLARIRAILRR-- 118
Db 61 LEVCRVRVKNVETPIIMWTARDSVIDRVSGLDHGADDDYIVKFPDIBELLARIRAILRRID 120

QY 119 -----OPQKDIIDVNGITIDKNAPKVTNGABEELTKTEYDLYLLAENKHNVMOREQI 172
Db 121 LESQSQSTKQTTTYDKLTIKENLVVKGDEVINLTREYELLTLTMENINVVLARDVL 180

QY 173 LNHVGNSEVETNVVDVYIRYLRNKLKPYDRDKMIETVRGVGVIR 219
Db 181 LNKVWGSEVETNVVDVYIRYLRNKLKIDRPGKSYIQTVRGTGVIR 227

RESULT 6
Q92BX8 PRELIMINARY; PRT; 226 AA.
AC Q92BX8
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Two-component response regulator.
GN LISR OR LMO1377 OR LIN1414.
OS Listeria monocytogenes, and
OS Listeria innocua.
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OX NCBI_TaxID=1639, 1642;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=L.monocytogenes, and L.innocua;
RC STRAIN=EGD-e / Serovar 1/2a, and CLIP 11262 / Serovar 6a;
RX MEDLINE=21537279; PubMed=11679669;
RA Glaser P., Frangeul P., Bloecher H., Brandt P., Chakraborty T.,
RA Baquero F., Berche P., Buchrieser C., Rusniok C., Amend A.,
RA Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA Entian K.-D., Feihi H., Garcia-del Portillo F., Garrido P.,
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA Jones L.-M., Kaerst U., Krest J., Kuhn M., Kunst F., Kurapkat G.,
RA Madoeno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
RA Nordsiek G., Novellia S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose W., Schlueter T., Simoes N., Tierrez A.,
RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;
RT "Comparative genomics of Listeria species.";
RL Science 294:849-852(2001).
DR EMBL; AL591978; CAC99455.1; -
DR EMBL; AL596168; CAC96645.1; -
DR PIR; AE1609; AE1609.
DR PIR; A11246; A11246.
DR ListiList; LMO1377; -
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0005215; F:two-component response regulator activity; IEA.
DR GO; GO:0000156; F:two-component response regulator activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR GO; GO:0000160; P:two-component signal transduction system (p. . .); IEA.
DR InterPro; IPR001789; Response_reg.
DR InterPro; IPR005829; Sug_transporter.
DR InterPro; IPR001867; Trans_reg_C.
DR Pfam; PF00072; response_reg; 1.
DR Pfam; PF00486; trans_reg_C; 1.
DR ProDom; PD000039; Response_reg; 1.
DR ProDom; PD000329; Trans_reg_C; 1.
DR SMART; SM00448; REC; 1.
DR PROSITE; PS00110; RESPONSE_REGULATORY; 1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW Complete proteome.
SQ SEQUENCE 226 AA; 26419 MW; BE2598E0254FF053 CRC64;

```

```

Query Match 56.0%; Score 628; DB 16; Length 226;
Best Local Similarity 55.1%; Pred. No. 5.9e-40;
Matches 124; Conservative 39; Mismatches 56; Indels 6; Gaps 1;

QY 1 MTQILIVEDQNLFARFLELELTHENVNVDTEYDQDGLDKALSHYVDLIILDLMLPSING 60
Db 1 MNRILIVEDKNLFARFLELELQHENYETAVANDGRAGLEALNEEMDAILLDLMLPLNG 60

QY 61 LEICRKIROOQSTPIIIITAKSDTYDKVAGLDYGADDDYIVKFPDIBELLARIRAILRR-- 118
Db 61 VEVCRRVRVQKTPITARDSVIDRVSGLDHGADDDYIVKFPDIBELLARIRAILRRVE 120

QY 119 -----OPQKDIIDVNGITIDKNAPKVTNGABEELTKTEYDLYLLAENKHNVMOREQI 174
Db 121 NAEQSAKQTTLYQYRNLIVKDEIRIVKDEIDLTREYELLTLTMENINVLTREVLN 180

QY 175 HVHGNSEVETNVVDVYIRYLRNKLKPYDRDKMIETVRGVGVIR 219
Db 181 KVMGYETETNVVDVYVYLRNKLIDHPDESVIQTVRGTGVYMR 225

RESULT 7
Q98B10 PRELIMINARY; PRT; 229 AA.
ID Q98B10
AC Q98B10
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Response regulators consisting of a CheY-like receiver domain and a
DE HTH DNA-binding domain.
GN OMPR3 OR TTE1016.
OS Thermoanaerobacter tengcongensis.
OC Bacteria; Firmicutes; Clostridia; Thermoanaerobacteriales;
OC Thermoanaerobacteriaceae; Thermoanaerobacter.
OX NCBI_TaxID=119072;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=M84 / JCM 11007;
RC MEDLINE=21992816; PubMed=11997336;
RA Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,
RA Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,
RA Tan H., Chen R., Wang J., Yu J., Yang H.;
RT "A complete sequence of T. tengcongensis genome.";
RL Genome Res 12:689-700(2002).
DR EMBL; AE013066; AM24271.1; -
DR HSSP; P06143; ID42.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0000156; F:two-component response regulator activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR001789; Response_reg.
DR InterPro; IPR005829; Sug_transporter.
DR InterPro; IPR001867; Trans_reg_C.
DR Pfam; PF00072; response_reg; 1.
DR Pfam; PF00486; trans_reg_C; 1.
DR ProDom; PD000039; Response_reg; 1.
DR ProDom; PD000329; Trans_reg_C; 1.
DR SMART; SM00448; REC; 1.
DR PROSITE; PS00110; RESPONSE_REGULATORY; 1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW Complete proteome.
SQ SEQUENCE 229 AA; 26618 MW; 1097D71DEF3A5ACD CRC64;

Query Match 52.5%; Score 588.5; DB 16; Length 229;
Best Local Similarity 53.4%; Pred. No. 5.9e-37;
Matches 118; Conservative 38; Mismatches 60; Indels 5; Gaps 4;

QY 3 QILIVEDQNLFARFLELELTHENVNVDTEYDQDGLDKALSHYVDLIILDLMLPSINGLE 62
Db 4 KILIVEDQRIARFLQLEFEGYEVKIVYDGTSGVKEALEGVDLVLLDMLPGMDGYE 63

```



```

Db      185 WGYEETNTVNDVYIRYLNRKIDVEQDSYIQTIRGLGVYMR 227
RESULT 10
Q9RPZ0 PRELIMINARY; PRT; 226 AA.
AC Q9RPZ0;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE LlsR.
GN LlsR.
OS Listeria monocytogenes.
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OX NCBI_TaxID=1639;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LO28;
RX MEDLINE=20011355; PubMed=10542190;
RA Cotter P.D., Emerson N., Gahan C.G., Hill C.;
RT "Identification and disruption of llsRK, a genetic locus encoding a
RT two-component signal transduction system involved in stress tolerance
RT and virulence in Listeria monocytogenes.";
RL J. Bacteriol. 181:6840-6843 (1999).
DR EMBL; AF139908; AAF03932.1; -.
DR HSP; P08402; I800
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0000156; F:two-component response regulator activity; IEA.
DR GO; GO:0007600; P:sensory perception; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0000160; P:two-component signal transduction system (p. . .; IEA.
DR InterPro; IPR001789; Response reg.
DR InterPro; IPR001867; Trans_reg_C.
DR Pfam; PF000072; response_reg; 1.
DR Pfam; PF00486; trans_reg_C; 1.
DR ProDom; PD0000329; Response reg; 1.
DR ProDom; PD000329; Trans_reg_C; 1.
DR SMART; SM00448; REC; 1.
DR PROSITE; PS00110; RESPONSE_REGULATORY; 1.
KW DNA-binding; Phosphorylation; Sensory transduction; Transcription;
KW Transcription regulation.
SQ SEQUENCE 226 AA; 26541 MW; 802DA0E0D1A26EFC CRC64;

Query Match 51.68; Score 579; DB 2; Length 226;
Best Local Similarity 52.2%; Pred. No. 3.1e-36;
Matches 120; Conservative 38; Mismatches 56; Indels 16; Gaps 3;

QY 1 MTQILIVDEQNARFLELELTHENVVDYDQGLDKALSHYDILDLMLPSING 60
Db 1 MNRILIVDEKNLARFLELELQHENYETAVNDGRAGLEALNEWDAILDLMLPHNG 60

QY 61 LEICRKIQOOSTPIIITAKSDPYDKVAGLDY-----GADDYIVKPFDELLARIRAI 115
Db 61 VEVCRVRVQKQTPILMITARD-----LLSIVYPMWITGADDYIVKPFDELLARLSL 115

QY 116 LRR-----QPKDIIIVNGITIDKNAFKYTVNGAEILTKTEYDILLYLAENKHMVQR 169
Db 116 LRRVENAEQSAKQTTLOYRLNIVEKENRIYKRDDEEIIIDLTKREYELLTLNENNVILTR 175

QY 170 EQILNHVWYNSVETNVVDYIRYLNRKLPYDRDKMIETVRGVGVIR 219
Db 176 EVLLNKVWGYETEVETNVVDYIRYLNRKIDHPDESYIQTIRGTGVYMR 225

RESULT 11
Q8E3T4 PRELIMINARY; PRT; 229 AA.
AC Q8E3T4;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN GBS1672.

OS Streptococcus agalactiae (serotype III).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=216495;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NEM316 / Serotype III;
RX MEDLINE=22242508; PubMed=12354221;
RA Glaser P., Rusniok C., Buchrieser C., Chevalier F., Frangeul L.,
RA Msadek T., Zouine M., Couve E., Lalioui L., Poyart C., Trieu-Cuot P.,
RA Kunst F.;
RT "Genome sequence of Streptococcus agalactiae, a pathogen causing
RT invasive neonatal disease.";
RL Mol. Microbiol. 45:1499-1513 (2002).
DR EMBL; AL768852; CAD47331.1; -.
DR Sagal1st; GBS1672; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0000156; F:two-component response regulator activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR GO; GO:0000160; P:two-component signal transduction system (p. . .; IEA.
DR InterPro; IPR001789; Response reg.
DR InterPro; IPR005829; Sug transporter.
DR InterPro; IPR001867; Trans_reg_C.
DR Pfam; PF000072; response_reg; 1.
DR Pfam; PF00486; trans_reg_C; 1.
DR ProDom; PD0000329; Response reg; 1.
DR ProDom; PD000329; Trans_reg_C; 1.
DR SMART; SM00448; REC; 1.
DR PROSITE; PS00110; RESPONSE_REGULATORY; 1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 229 AA; 26547 MW; 8830F539372AE2D8 CRC64;

Query Match 51.2%; Score 574.5; DB 16; Length 229;
Best Local Similarity 53.3%; Pred. No. 6.8e-36;
Matches 121; Conservative 38; Mismatches 55; Indels 13; Gaps 3;

QY 3 QILIVDEQNARFLELELTHENVVDYDQGLDKALSHYDILDLMLPSINGLE 62
Db 4 KILIIEDKNLARFVSLLELHEGYDVVWETNGREGIDTALEKDFDILDLMLPEMDGPE 63

QY 63 ICRKIQOOSTPIIITAKSDPYDKVAGLDYDADDYIVKPFDELLARIRAI LRQPOK 122
Db 64 ITRLQAEKTYIMMTARDSDVNDIVAGLDRGADDYIVKPFDELLARVLRPRQIE 123

QY 123 DIIDVNGITIDKNAFK-VTVN-----GAEILTKTEYDILLYLAENKHMVQREI 172
Db 124 TKTKEG---DSGSFRDLSLNTHNSAMRGDEISLTKEFDLLNVLMNVRVMTREEL 180

QY 173 LNHVWYNSVETNVVDYIRYLNRKLPYDRDKMIETVRGVGVIR 219
Db 181 LEHVWKYDVAETNVVDYIRYLNRKIDIPGRESYIQTIRGVGVIR 227

RESULT 12
Q8DY68 PRELIMINARY; PRT; 229 AA.
AC Q8DY68;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE DNA-binding response regulator CsrR.
GN CSRR OR SAG1625.
OS Streptococcus agalactiae (serotype V).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=216466;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=2603 V/R / Serotype V;
RC MEDLINE=2222998; PubMed=12200547;

```



```

OX NCBI_TaxID=1309;
RN [1];
RP SEQUENCE FROM N.A.
RC STRAIN=109CS;
RA MEDLINE=20263759; PubMed=10802169;
RA Sato Y., Yamamoto Y., Kizaki H.;
RT "Construction of partial duplication mutants (merodiploid mutants) to
RT investigate function of genes in vivo: identification of a regulatory
RT gene for the glucan-binding protein C of Streptococcus mutans."
RL FEMS Microbiol. Lett. 186:187-191(2000).
RN [2];
RP SEQUENCE FROM N.A.
RC STRAIN=UA159 / ATCC 700610 / Serotype C;
RX MEDLINE=22295063; PubMed=12397186;
RA Ajdic D., McShan W.M., McLaughlin R.E., Savic G., Chang J.,
RA Carson M.B., Primeaux C., Tian R., Kenton S., Jia H., Lin S., Qian Y.,
RA Li S., Zhu H., Najjar F., Lai H., White J., Roe B.A., Ferretti J.J.;
RT "Genome sequence of Streptococcus mutans UA159, a cariogenic dental
RT pathogen."
RL Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439(2002).
DR ENBL; AB024522; BA85986.1; -.
DR ENBL; AE015017; AAN59535.1; -.
DR HSSP; P08402; 1B00.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0000156; F:two-component response regulator activity; IEA.
DR GO; GO:0007600; P:sensory perception; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR GO; GO:0001600; P:two-component signal transduction system (p. . .; IEA.
DR InterPro; IPR001789; Response_reg.
DR InterPro; IPR005829; Sug_transporter.
DR InterPro; IPR001867; Trans_reg_C.
DR Pfam; PF00486; trans_reg_C; 1.
DR ProDom; PD000329; Response_reg; 1.
DR SMART; SM00448; REC; 1.
DR PROSITE; PS00110; RESPONSE_REGULATORY; 1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
DR DNA-binding; Phosphorylation; Sensory transduction; Transcription;
KW Transcription regulation; Complete proteome.
SQ SEQUENCE 230 AA; 26739 MW; 71A570DC5AD48B20 CRC64;

Query Match 50.8%; Score 570.5; DB 16; Length 230;
Best Local Similarity 52.0%; Pred. No. 1.4e-35;
Matches 176; Conservative 36; Mismatches 64; Indels 7; Gaps 1;

QY 4 ILIVDEQNARLELELTHENVNVDTEYDQGLDKALSHYVDLIILDLMLPSINGLEI 63
DB 5 ILIIEDEKLAKFVSLEQHEGYEVETVDGRSGLERALEKDFDILLDLMLPMDGYEV 64
QY 64 CRKIROQOSTPIIIITAKSDTYDKVAGLDYGDADYIVKPFDEIELLARIRLRQ--- 119
DB 65 IRLQLEKTYVITITARDSIMIVTGLDRGADGYIVKPFDEIELLARIRLRQDIEK 124
QY 120 ---PKQDIIDVNGITIDKNAPKVTNGAEIELTKTEYDLYLLAENKNHVMOREQILN 176
DB 125 AKHSSKNDTTYRDLKLVHNRAIRNGETIPLTKREFDNLNVALVENINQVMTREELLARV 184
QY 177 WGYNSEVETNVVDYIRYLKPKYDRDKMIETVRGVGVIR 219
DB 185 WKYDQEAETNVVDYIRYLKPKYDRDKMIETVRGVGVIR 227

RESULT 15
Q8P2J8 PRELIMINARY; PRT; 228 AA.
AC Q8P2J8;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

```

```

DE COVR.
GN COVR OR SPYM18 0328.
OS Streptococcus Pyogenes (serotype M18).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=186103;
RN [1];
RP SEQUENCE FROM N.A.
RC STRAIN=NGAS8232 / Serotype M18;
RX MEDLINE=21927593; PubMed=11917108;
RA Smoot J.C., Barbican K.D., Van Gompel J.J., Smoot L.M., Chaussee M.S.,
RA Sylva G.L., Sturdevant D.E., Ricklefs S.M., Porcella S.F.,
RA Parkins L.D., Beres S.B., Campbell D.S., Smith T.M., Zhang Q.,
RA Kapur V., Daly J.A., Veasy L.G., Musser J.M.;
RT "Genome sequence and comparative microarray analysis of serotype M18
RT group A Streptococcus strains associated with acute rheumatic fever
RT outbreaks."
RL Proc. Natl. Acad. Sci. U.S.A. 99:4668-4673(2002).
DR ENBL; AS009978; AAL97083.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0000156; F:two-component response regulator activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR GO; GO:0001600; P:two-component signal transduction system (p. . .; IEA.
DR InterPro; IPR001789; Response_reg.
DR InterPro; IPR005829; Sug_transporter.
DR InterPro; IPR001867; Trans_reg_C.
DR Pfam; PF00072; response_reg; 1.
DR Pfam; PF00486; trans_reg_C; 1.
DR ProDom; PD000329; Response_reg; 1.
DR PROSITE; PS00110; RESPONSE_REGULATORY; 1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW Complete proteome.
SQ SEQUENCE 228 AA; 26616 MW; 34C24A5DB6A35113 CRC64;

Query Match 50.6%; Score 568; DB 16; Length 228;
Best Local Similarity 52.4%; Pred. No. 2.1e-35;
Matches 118; Conservative 38; Mismatches 59; Indels 10; Gaps 2;

QY 3 QILIVDEQNARLELELTHENVNVDTEYDQGLDKALSHYVDLIILDLMLPSINGLE 62
DB 4 KILLIIEDEKNLARFVSLEQHEGYEVIVEVNGEGLETALEKEFDLILLDLMLPMDGFE 63
QY 63 ICRKIROQOSTPIIIITAKSDTYDKVAGLDYGDADYIVKPFDEIELLARIRLRQ--- 119
DB 64 VTRRLQTEKTYVITIMMTARDSIMDVVAGLDRGADYIVKPFDEIELLARIRLRQDIE 123
QY 120 ---PKQDIIDVNGITIDKNAPKVTNGAEIELTKTEYDLYLLAENKNHVMOREQILN 174
DB 124 SEKKVPSQGI--YRDLVLPQNRSVNRGDDEIPLTKREYDNLNLTNNMNRVMTRELLS 181
QY 175 HWGYNSEVETNVVDYIRYLKPKYDRDKMIETVRGVGVIR 219
DB 182 NWKYDEAVETNVVDYIRYLKPKYDRDKMIETVRGVGVIR 226

Search completed: October 4, 2004, 18:09:31
Job time : 52.3761 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 5, 2004, 00:16:04 ; Search time 2948.76 Seconds
(without alignments)
10818.262 Million cell updates/sec

Title: US-09-006-627-3
Perfect score: 736
Sequence: 1 ATTTACGTTTGTGTCATCGTA.....TTAAATTTGTATTTTCTA 736

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues
Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb.ba.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.on.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sts.*
- 12: gb.sy.*
- 13: gb.un.*
- 14: gb.vi.*
- 15: em.ba.*
- 16: em.fun.*
- 17: em.hum.*
- 18: em.in.*
- 19: em.mu.*
- 20: em.om.*
- 21: em.or.*
- 22: em.ov.*
- 23: em.pat.*
- 24: em.ph.*
- 25: em.pl.*
- 26: em.ro.*
- 27: em.sts.*
- 28: em.un.*
- 29: em.vi.*
- 30: em.htg.hum.*
- 31: em.htg.inv.*
- 32: em.htg.other.*
- 33: em.htg.mus.*
- 34: em.htg.pln.*
- 35: em.htg.rod.*
- 36: em.htg.mam.*
- 37: em.htg.vrt.*
- 38: em.sv.*
- 39: em.htgo.hum.*
- 40: em.htgo.mus.*
- 41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	736	100.0	736	6	A63826	A63826 Sequence 1
2	736	100.0	736	6	AR306347	AR306347 Sequence
3	736	100.0	736	6	BD136061	BD136061 Histidine
C 4	736	100.0	2201	6	AR306346	AR306346 Sequence
C 5	736	100.0	2201	6	BD136060	BD136060 Histidine
C 6	736	100.0	3083	1	AF165314	AF165314 Staphyloc
7	736	100.0	346900	1	AP003362	AP003362 Staphyloc
C 8	726	98.6	1440	6	AR354607	AR354607 Sequence
9	723.2	98.3	295350	1	AP004826	AP004826 Staphyloc
10	694	94.3	303750	1	AP003133	AP003133 Staphyloc
C 11	657	89.3	657	6	AX621578	AX621578 Sequence
12	412.8	56.1	300892	1	AE016747	AE016747 Staphyloc
13	406.4	55.2	3099	1	AF269746	AF269746 Staphyloc
14	406.4	55.2	3099	6	AX145064	AX145064 Sequence
C 15	390.4	53.0	3230	1	AF269630	AF269630 Staphyloc
C 16	390.4	53.0	3230	6	AX144950	AX144950 Sequence
17	347	47.1	3018	1	AF270244	AF270244 Staphyloc
18	347	47.1	3018	6	AX145562	AX145562 Sequence
C 19	302.6	41.1	501	6	AX142701	AX142701 Sequence
C 20	302.6	41.1	501	6	AX144065	AX144065 Sequence
21	224.8	30.5	195269	6	AX417035	AX417035 Sequence
C 22	224.8	30.5	333050	1	AL596168	AL596168 Listeria
23	224.8	30.5	349980	6	AX413017	AX413017 Sequence
C 24	224.8	30.5	349980	6	AX417044	AX417044 Sequence
C 25	204.2	27.7	301488	1	AE016950	AE016950 Enterococ
C 26	202.2	27.5	250050	1	AL591978	AL591978 Listeria
C 27	202.2	27.5	349980	6	AX641668	AX641668 Sequence
28	193.6	26.3	10575	1	AE006390	AE006390 Lactococc
29	193.6	26.3	20601	1	AE014265	AE014265 Streptoco
30	193.6	26.3	44145	6	AX602195	AX602195 Sequence
31	193.6	26.3	174050	1	SAG766852	AL766852 Streptoco
C 32	193.4	26.3	690	6	AX607121	AX607121 Sequence
33	193	26.2	1908	6	BD193433	BD193433 Enterococ
C 34	187.8	25.5	3619	1	AF139908	AF139908 Listeria
C 35	187.2	25.4	687	6	AR348271	AR348271 Sequence
C 36	187.2	25.4	5001	1	LLU81166	U81166 Lactococcus
C 37	185.4	25.2	1053	1	AB024522	AB024522 Streptoco
C 38	185.4	25.2	12451	1	AE015017	AE015017 Streptoco
C 39	184.6	25.1	302050	1	AL935256	AL935256 Lactobaci
40	181.6	24.7	11726	1	AE010631	AE010631 Fusobacte
C 41	181	24.6	10816	1	AE013066	AE013066 Thermoana
C 42	179.4	24.4	10894	1	AE010570	AE010570 Fusobacte
C 43	175	23.8	12434	1	AE006498	AE006498 Streptoco
C 44	174.8	23.8	11581	1	AE007581	AE007581 Clostridi
C 45	173.4	23.6	2472	1	AF082668	AF082668 Streptoco

ALIGNMENTS

RESULT 1
A63826
LOCUS A63826 736 bp DNA linear PAT 12-MAR-1998
DEFINITION Sequence 1 from Patent WO9723506.
ACCESSION A63826
VERSION A63826.1 GI:3717394
KEYWORDS .
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1
Wallis,N. and Hodgson,J.E.
TWO-COMPONENT SIGNAL TRANSDUCING SYSTEM FROM STAPHYLOCOCCUS AUREUS
Patent: WO 9723506-A 1 03-JUL-1997;
JOURNAL SMITHKLINE BEECHAM PLC (GB)

[illegible]


```

ACCESSION   BD136061
VERSION     BD136061.1  GI:23231006
KEYWORDS    JP 2002508955-A/2.
SOURCE      Staphylococcus aureus
ORGANISM    Staphylococcus aureus
REFERENCE   Bacteria; Firmicutes; Bacillales; Staphylococcus.
AUTHORS     Wallis,N.G., Schilling,L.K., Mooney,J.L., Debouck,C., Twon,I.I.,
            Jaworski,D., Wang,M. and Throup,J.P.
TITLE       Histidine kinase
JOURNAL     Patent: JP 2002508955-A 2 26-MAR-2002;
            SMITHKLINE BEECHAM CORP
COMMENT     OS Staphylococcus aureus
            PN JP 2002508955-A/2
            PD 26-MAR-2002
            PF 12-JAN-1999 JP 2000540212
            PR 13-JAN-1998 US 09/006627
            PI NICOLA G WALLIS,LISA K SCHILLING,JEFFREY L MOONEY,CHRISTINE
            DEBOUCK,
            FI I I TWON,DEBORAH D JAWORSKI,MING WANG,JOHN P THROUP PC
            C12N15/09,A61K38/00,A61K38/51,A61K39/02,A61K39/395,A61K39/395, PC
            A61K45/00.
            PC A61P31/04,A61P37/04,C07H21/02,C07K16/40,C12N1/21,C12N5/10, PC
            C12N9/12,C12Q1/48,C12Q1/68,C12N15/00,A61K37/02,A61K37/56,C12N5/00 CC
            Histidine kinase
            FH Key Location/Qualifiers
            FT source 1..736
            FT /organism='Staphylococcus aureus'.
FEATURES    source
            LOCATION 1..736
            /organism='Staphylococcus aureus'
            /mol_type='genomic DNA'
            /db_xref='taxon:1280'
ORIGIN
Query Match 100.0%; Score 736; DB 6; Length 736;
Best Local Similarity 100.0%; Pred. No. 9.3e-125;
Matches 736; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATTTACGTTTGTGTCATGATACATACCCAGCCAGCACTGTTTCAATCATTTTGTGTC 60
Db 1 ATTTACGTTTGTGTCATGATACATACCCAGCCAGCACTGTTTCAATCATTTTGTGTC 60
Qy 61 ACGATCGTATGTTTAACTGTTTTCGTAATAATCTTATATAAATCTTATATAAATCTGACATTTGT 120
Db 61 ACGATCGTATGTTTAACTGTTTTCGTAATAATCTTATATAAATCTGACATTTGT 120
Qy 121 TTCTACTTCACTATATATACCCCATACATGATTTAAATTTGTTCCCGTTGCATAACATG 180
Db 121 TTCTACTTCACTATATATACCCCATACATGATTTAAATTTGTTCCCGTTGCATAACATG 180
Qy 181 GTTTTATTATTTTCAGTACAGATATAGTAATCATCTCTGTTTGTGTTTCAATTTTC 240
Db 181 GTTTTATTATTTTCAGTACAGATATAGTAATCATCTCTGTTTGTGTTTCAATTTTC 240
Qy 241 TGGCGCATTTACCGTCACCTTTTAAAGCGTTCTTATCAATTTGTAATACCGTTGACATCGAT 300
Db 241 TGGCGCATTTACCGTCACCTTTTAAAGCGTTCTTATCAATTTGTAATACCGTTGACATCGAT 300
Qy 301 AATATCCCTTTTGTGGCTGACGACGTAAATTTGCAGCAATTTGCTAAAAGTTCTTCAAT 360
Db 301 AATATCCCTTTTGTGGCTGACGACGTAAATTTGCAGCAATTTGCTAAAAGTTCTTCAAT 360
Qy 361 ATCAACGGCTTAACATATATATGTCGTGCAACCGTAATCAAGCCAGCACTTTGTGTCATA 420
Db 361 ATCAACGGCTTAACATATATATGTCGTGCAACCGTAATCAAGCCAGCACTTTGTGTCATA 420
Qy 421 CGTATCACTTTTCCGTGTAATTTAATATGATAGGTGATGTTGTTGCTCTAAATTTTGGC 480
Db 421 CGTATCACTTTTCCGTGTAATTTAATATGATAGGTGATGTTGTTGCTCTAAATTTTGGC 480
Qy 481 ACAAATTTCTAAGCCATTTAATTTGACGGCAACATTTAAATCTAATATGATTTAAATCATAGTA 540

```

```

Db 481 ACAAATTTCTAAGCCATTTAATTTGACGGCAACATTTAAATCTAATATGATTTAAATCATAGTA 540
Qy 541 ATGGCTAAGCGCTTTATCTAAACCGTCTTGTGTCATCTGTCGTCACATTTGTAAT 600
Db 541 ATGGCTAAGCGCTTTATCTAAACCGTCTTGTGTCATCTGTCGTCACATTTGTAAT 600
Qy 601 TTCTATGTCAGTTTCCCAATTCAGAAATCTTGTGTAAGTTTGTTCATCTTCTACTATTA 660
Db 601 TTCTATGTCAGTTTCCCAATTCAGAAATCTTGTGTAAGTTTGTTCATCTTCTACTATTA 660
Qy 661 AATTTGGCTCATTTTGTACACCTCATATACACCTTTTCTAATAGGTAATATATTTTAA 720
Db 661 AATTTGGCTCATTTTGTACACCTCATATACACCTTTTCTAATAGGTAATATATTTTAA 720
Qy 721 ATTTTGTATTTTCTTA 736
Db 721 ATTTTGTATTTTCTTA 736
RESULT 4
LOCUS       AR306346 2201 bp DNA linear PAT 12-JUN-2003
DEFINITION Sequence 1 from patent US 6548281.
ACCESSION   AR306346
VERSION     AR306346.1 GI:31696114
KEYWORDS    Unknown.
SOURCE      Unknown.
ORGANISM    Unclassified.
REFERENCE   1 (bases 1 to 2201)
AUTHORS     O'Dwyer,K.M., Jaworski,D.D., Mooney,J.L., Shilling,L.K.,
            Throup,J.P., Wallis,N.G., Zhong,Y.Y. and Wang,M.
TITLE       Histidine kinase
JOURNAL     Patent: US 6548281-A 1 15-APR-2003;
FEATURES    Location/Qualifiers
            source 1..2201
            /organism='unknown'
            /mol_type='genomic DNA'
ORIGIN
Query Match 100.0%; Score 736; DB 6; Length 2201;
Best Local Similarity 100.0%; Pred. No. 7.5e-125;
Matches 736; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATTTACGTTTGTGTCATGATACATACCCAGCCAGCACTGTTTCAATCATTTTGTGTC 60
Db 759 ATTTACGTTTGTGTCATGATACATACCCAGCCAGCACTGTTTCAATCATTTTGTGTC 700
Qy 61 ACGATCGTATGTTTAACTGTTTTCGTAATAATCTTATATAAATCTTATATAAATCTGACATTTGT 120
Db 699 ACGATCGTATGTTTAACTGTTTTCGTAATAATCTTATATAAATCTGACATTTGT 640
Qy 121 TTCTACTTCACTATATATACCCCATACATGATTTAAATTTGTTCCCGTTGCATAACATG 180
Db 639 TTCTACTTCACTATATATACCCCATACATGATTTAAATTTGTTCCCGTTGCATAACATG 580
Qy 181 GTTTTATTATTTTCAGTACAGATATAGTAATCATCTCTGTTTGTGTTTCAATTTTC 240
Db 579 GTTTTATTATTTTCAGTACAGATATAGTAATCATCTCTGTTTGTGTTTCAATTTTC 520
Qy 241 TGGCGCATTTACCGTCACCTTTTAAAGCGTTCTTATCAATTTGTAATACCGTTGACATCGAT 300
Db 519 TGGCGCATTTACCGTCACCTTTTAAAGCGTTCTTATCAATTTGTAATACCGTTGACATCGAT 460
Qy 301 AATATCCCTTTTGTGGCTGACGACGTAAATTTGCAGCAATTTGCTAAAAGTTCTTCAAT 360
Db 459 AATATCCCTTTTGTGGCTGACGACGTAAATTTGCAGCAATTTGCTAAAAGTTCTTCAAT 400
Qy 361 ATCAACGGCTTAACATATATGTCGTGCAACCGTAATCAAGCCAGCACTTTGTGTCATA 420
Db 399 ATCAACGGCTTAACATATATGTCGTGCAACCGTAATCAAGCCAGCACTTTGTGTCATA 340

```



```
FEATURES
source
Location/Qualifiers
1..3083
/organism="Staphylococcus aureus"
/mol_type="genomic DNA"
/strain="ISP794"
/db_xref="taxon:1280"
336..995
/gene="arlR"
336..995
/gene="arlR"
/codon_start=1
/transl_table=11
/product="putative response regulator ArlR"
/protein_id="AAF85896.1"
/db_xref="GI:9230552"
/translation="MTQILIVDEQNLARFLELELTHENVNVDTEYDQGLDKALSH
YDILILMLPSNGICRIRQOOSTRIIITAKSDTYDKVAGLDYGDDYIVKP
FDIELLARIARLRQKQIIDVNGITIDKNAPKVTVNGAEELTKETFDLYLLA
ENKHMVQREQLLNHWGYNSEVETNVVDVYIRLNRKLPYDRDKMIETVRGVGYVI
R"
992..2347
/gene="arlS"
992..2347
/gene="arlS"
/codon_start=1
/transl_table=11
/product="putative protein histidine kinase ArlS"
/protein_id="AAF85897.1"
/db_xref="GI:9230553"
/translation="MTRKLRNNWIIVTMTFTVIFLCLIIFFLKDTHLSELDD
AERSSDINNLFHSPKVDIISALDNLASLGNFQIIIIYDEHNNKLFETSDNTVRVEP
AEYHRYDFRIVKRYKGYEYLIIEKPIITQDFKGYSLLIHSLNVDNIKSLYIIALA
FGVATITATISVVFSTQITKPLVSLNKEIRRDGNFQKLQNTNRYEIDLNANT
FNEWMSQIEFSFNQORVEDASHELETPLOIIOGHNLIOGWKKOPAVLEESINLS
IEENWRIKLVEELLELPKDVNDISSEAQTVDHINDEIRSLHSLKQLHPYQEDTDL
TSKLEIKKPKHQFEQLPFIIDNAIKYDVNRNKKIKVTRLNKKQIKIIBITDGHGIP
BEDQDFIPDRFYRVDKRSRSGGNGLSIAQKIIQLNGSGSIKIKSEINKGTTFKII
F"
Query Match 100.0%; Score 736; DB 1; Length 3083;
Best Local Similarity 100.0%; Pred. No. 7.le-135;
Matches 736; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATTTACGTTTGTCTATCATACCAACGCGCAGCACTGTTTCAATCATTTTGTG 60
DB 1007 ATTTACGTTTGTCTATCATACCAACGCGCAGCACTGTTTCAATCATTTTGTG 948
QY 61 ACGATCGTATGTTTAACTGTTTTCGTAATAATCTTATATAAACAATCTACGACATTTGT 120
DB 947 ACGATCGTATGTTTAACTGTTTTCGTAATAATCTTATATAAACAATCTACGACATTTGT 888
QY 121 TTCTACTTCACTATTATACCCCAATGATTTAAATTTGTTCCCGTTGCAATCAATG 180
DB 887 TTCTACTTCACTATTATACCCCAATGATTTAAATTTGTTCCCGTTGCAATCAATG 828
QY 181 GTTTTATTTTTCAGCTAGAGATAGTAATCATCTGTTTGTGTTTGAATTCATTTTC 240
DB 827 GTTTTATTTTTCAGCTAGAGATAGTAATCATCTGTTTGTGTTTGAATTCATTTTC 768
QY 241 TCGCGCATTTACCGTCACCTTTAAAGCGTCTTATCAATTTGTAATACCGTTGACATCGAT 300
DB 767 TCGCGCATTTACCGTCACCTTTAAAGCGTCTTATCAATTTGTAATACCGTTGACATCGAT 708
QY 301 AATATCCTTTTGTGCTGACGCGTAAATTTGCAAGAAATCTTGTGTAAGATTTCTCAAT 360
DB 707 AATATCCTTTTGTGCTGACGCGTAAATTTGCAAGAAATCTTGTGTAAGATTTCTCAAT 648
QY 361 ATCAACCGCTTAACATATATATCTGCGACCGTAATCAAGCCAGCACTTTGTGCTATA 420
DB 647 ATCAACCGCTTAACATATATATCTGCGACCGTAATCAAGCCAGCACTTTGTGCTATA 588
QY 421 CGTATCACTTTTCGCTGTAATATATAATGATAGGTGTAGATTGTTGTTCTAAATTTTGGC 480
```

```

587 CGTATCACTTTTCGCTGTAATATATAATGATAGGTGTAGATTGTTGTTCTAAATTTTGGC 528
481 ACAAATTTCTAAGCCATTAAATTGACGCAACATTAAATCTAATATGATTAATCATAGTA 540
527 ACAAATTTCTAAGCCATTAAATTGACGCAACATTAAATCTAATATGATTAATCATAGTA 468
541 ATGGCTAAGCGCTTTATCTAAACCGTCTTGTCCATCATCTCTGTGTCACCATTTGTAAT 600
467 ATGGCTAAGCGCTTTATCTAAACCGTCTTGTCCATCATCTCTGTGTCACCATTTGTAAT 408
601 TTCAATGTGAGTTCCCAATTCAGAAATCTTGTCAAGTTTGTTCATCTTCTACTATTAA 660
407 TTCAATGTGAGTTCCCAATTCAGAAATCTTGTCAAGTTTGTTCATCTTCTACTATTAA 348
661 AATTGGCTCATTTGTACACCTCATATTACGACATTTTCTAATAAGGTAATATTTTAA 720
347 AATTGGCTCATTTGTACACCTCATATTACGACATTTTCTAATAAGGTAATATTTTAA 288
721 ATTTGTATTTTCTA 736
287 ATTTGTATTTTCTA 272

RESULT 7
LOCUS AP003362
DEFINITION Staphylococcus aureus subsp. aureus Mu50 genomic DNA, complete
sequence, section 5/9.
ACCESSION AP003362 BA000017
VERSION AP003362.2 GI:14247083
KEYWORDS
SOURCE
ORGANISM Staphylococcus aureus subsp. aureus Mu50
Bacteria; Firmicutes; Bacillales; Staphylococcus.
REFERENCE
1 Kuroda,M., Ohta,T., Uchiyama,I., Baba,T., Yuzawa,H., Kobayashi,I.,
Cui,L., Oguchi,A., Aoki,K., Nagai,Y., Lian,J., Ito,T., Kanamori,M.,
Matsumaru,H., Maruyama,A., Murakami,H., Hosoyama,A.,
Mizutani-Ui,Y., Takahashi,N.K., Sawano,T., Inoue,R., Kaito,C.,
Sekimizu,K., Hirakawa,H., Kuhara,S., Goto,S., Yabuzaki,J.,
Kanehisa,M., Yamashita,A., Oshima,K., Furuya,K., Yoshino,C.,
Shiba,T., Hattori,M., Ogasawara,N., Hayashi,H. and Hiramatsu,K.
Whole genome sequencing of methicillin-resistant Staphylococcus
aureus
Lancet 357 (9264), 1225-1240 (2001)
2 (bases 1 to 346900)
Direct Submission
Submitted (28-FEB-2001) Toshiko Ohta, University of Tsukuba College
of Medical Technology and Nursing, Department of Medical
Technology; 1-1-1 Ten-nodai, Tsukuba, Ibaraki 305-8577, Japan
(E-mail:tohta@sakura.cc.tsukuba.ac.jp, Tel:81-298-53-3454,
Fax:81-298-53-3454)
On May 29, 2001 this sequence version replaced gi:13875626.
COMMENT
Location/Qualifiers
1..346900
/organism="Staphylococcus aureus subsp. aureus Mu50"
/mol_type="genomic DNA"
/strain="Mu50"
/sub_species="aureus Mu50"
/db_xref="taxon:158878"
12..263
/gene="SAV1313"
12..263
/gene="SAV1313"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAB57475.1"
/db_xref="GI:14247084"
```

```
gene /translation="MLTIKLEKEQVIVYIPEDDISNGKGSVTFNKKDAEVIDFS
987. .1181 /gene="SAV1314"
CDS /gene="SAV1314"
987. .1181 /codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAB57476.1"
/db_xref="GI:14247089"
/translation="MNIQSYIGKLVKLTITNNKILIGKVIDFDKKVDFGYSNIEID
TGRITYDISENKITIVLQK"
gene complement (1725. .2750)
CDS /gene="SAV1315"
complement (1725. .2750)
/gene="SAV1315"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAB57477.1"
/db_xref="GI:14247086"
/translation="MISFENDYLEGAHEKVLKRLVDNLVQASGYGDFDQTAQAIK
KTDICPNKATRSADVETQNVVINSMEISYEGVISADTGHAVHGGAGIYSGHKV
ITIPSEKGVASDVETMETFKSDFKDHVPFGMVYISHPTEYGLTYSKSELEIC
KVCKYQLPFLMDGARLGYLMSQSDMTIKDIYAKYCDVFYIGTKIGALCKGDAI VFT
KNNEPKQFTTIRIKHGGALLAKRLTGIGLFTFDNLNLYENISRHALEMANKMGDGFN
KGYRLYFDSPTNQFFILSNKIAELQKVFAYWEKYDDQHRVRFATSWATTEENL
NKLELLI"
gene 3020. .3217
CDS /gene="SAV1316"
3020. .3217 /gene="SAV1316"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAB57478.1"
/db_xref="GI:14247087"
/translation="MEQIKLKTFTTAETLELEKNINAFLSSEBATNLKLVNITIK
ERTFPNNEEFNAILILSVNK"
gene 3238. .4755
CDS /gene="SAV1317"
3238. .4755 /gene="SAV1317"
/codon_start=1
/transl_table=11
/product="cardiolipin synthetase homolog"
/protein_id="BAB57479.1"
/db_xref="GI:14247088"
/translation="MFFVPPFIGEKTMRVTFSSNDLGLTFTIILAIGFIINLVAFI
FLERNRTASGAWLFLVFLPLIGFLVLPFRVTSARKKNKNGVLTDFDGLLK
QOIESPDKNYGTNDKQKHDLVSLMDQDGLTENNKVDFIDGNDLYDQVLKD
IKNAKEYIHLEYTFALDGLKRLHAEKLGQLEVKLYDDVGSKNVKNANFDHF
KSLGVEAEFAFKLPULNFRMNNRHKIIVIDGQLGVYGGNIDGYLGLGKLYW
RDLHLRGDVALQRLFDLWNSQAHREPOFQDYKFPKNGPLGNSPIQIAASGP
ASDHWQIEYQYKIMSAXKSVYLSQSPYFPDNSYINAIKAAKSGVDVHLMPCKPD
HPLVYATFNSADLLSSGVKIYTYENGFIHSKMCILDEIVSVGTANWDFPSFELN
EVNAFYVDENLXDLRVAYEHDITKSKQLTKESYANRPLSVKPKSLAKLVSPIL"
9939. .5436 /gene="SAV1318"
4939. .5436 /gene="SAV1318"
/notes="ATP-binding protein"
/codon_start=1
/transl_table=11
/product="ABC transporter homolog"
/protein_id="BAB57480.1"
/db_xref="GI:14247089"
/translation="MTQISNINKSFNRCVKNISPDIEGKCIALIGKNGAKSTLI
DILIGVWNSGRIFKDKLLOSENSIMEQKTMFPDQLKVIENLYOSFYNPFLPL
FEIILTKPDSSQLNQFNKLSGQQRLLDFVLSLIGQQLILLDEPTSTMDIEIKRY
FWSII"
5461. .5838 /gene="SAV1319"
5461. .5838 /gene="SAV1319"
/notes="ATP-binding protein"
/codon_start=1
/transl_table=11
/product="ABC transporter homolog"
/protein_id="BAB57481.1"
/db_xref="GI:14247090"
/translation="MLTSHYIEVERMSDKIILNENGEILLNDSTSHIRTNQSQIT
LSDEYIRKLUKDDDLVIQKNHNGTIKIITSNVNDTILYLOQLHINLDDIEIQKVSIV
DSYFNKKQKSGSYDTKLEENRI"
5807. .6538 /gene="SAV1320"
5807. .6538 /gene="SAV1320"
/notes="similar to ABC transporter integral membrane
protein"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAB57482.1"
/db_xref="GI:14247091"
/translation="MILSYLKIEFKVIMRKKTILSILFPVIFVILFTSILPEDEV
KPKFYKEYMYSMTVYSLLSGLLTFPLDIINEKQNEWRORLWVTPFTFSYVLSKVVK
TMLQPAIATLIVFMVGHFYKGVAMSAVQLESGIFLWLGASLLITFGILFSLANDIOK
TSALANIVTIGLAVGLNFPNTFPNWLQHVAVLPSYHLRKLGVDIASNNHINLIS
FAILLVUGSIIAVYCISHFKRAE"
6342. .7633 /gene="SAV1321"
6342. .7633 /gene="SAV1321"
/notes="similar to two-component sensor histidine kinase"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAB57483.1"
/db_xref="GI:14247092"
/translation="MKFLKDTSAIESLILYLPFIAGIFFNEVYCPKWLIIISVIF
SUSILVIVNNRLNLMFYILLIHYFICYFVSVHPMLSLFFYSAPAVPFTFN
NVKKTATNLFILMTICTIITLYLNNYFVAMVVYVVISLIMLDFKMKNNREYQKE
IAENRHNINTLAEQHRHIGQDLHTLGHFASLSKELAYKLIDTDVYKVAELL
AINKLSRESLNKVRRIIDDKLPSFTEIDSRKVKLDADIDFTFENKELAQLSPTK
QSMVMITREATNNVIKHAASKVHGKLTANNHKLLEDDGKIDSDCEVKSISQ
RVQHLNGTLAVDSTNGTKIILIEISTGGIA"
7630. .8232 /gene="SAV1322"
7630. .8232 /gene="SAV1322"
/notes="similar to two-component response regulator"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAB57484.1"
/db_xref="GI:14247093"
/translation="MTSLIIAEDQNMLOAMVOLIKLHGDGFELADTNDGLDAMKLE
EYNPNVILDIEMPGNTGLEAEIRKRLNIIKVIITFKPGYFEKAVVNDVDAIV
LKERSIEELVETINKVNGKEKYSATLMTSFFVDKNPLTPKQIIVLREITNGUSSKEI
SKKLFLTDGTVRNYTSVIIIDKLFDADRFDANKKANEKGWI"
complement (8345. .8533)
/gene="SAV1323"
complement (8345. .8533)
/gene="SAV1323"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAB57485.1"
/db_xref="GI:14247094"
/translation="WNISNENTVTLIAVIAIIIGIFLOIFKPLPLIVTAVLSILGI
FVGFIYILVYSFNKRKN"
8672. .9205 /gene="nuc"
8672. .9205 /gene="nuc"
8672. .9205
```


DEFINITION Staphylococcus aureus subsp. aureus MW2 DNA, complete genome,
 strain:MW2, section 5/10.
ACCESSION AF004826 BAO00033
VERSION AP004826.1 GI:21204263
KEYWORDS
SOURCE Staphylococcus aureus subsp. aureus MW2
ORGANISM Staphylococcus aureus subsp. aureus MW2
REFERENCE 1
AUTHORS Baba, T., Takeuchi, F., Kuroda, M., Yuzawa, H., Aoki, K., Oguchi, A., Nagai, Y., Iwama, N., Asano, K., Naimi, T., Kuroda, H., Cui, L., Yamamoto, K. and Hiramatsu, K.
TITLE Genome and virulence determinants of high virulence community-acquired MRSA
JOURNAL Lancet 359 (9320), 1819-1827 (2002)
MEDLINE 22040717
PUBMED 12044378
REFERENCE 2 (bases 1 to 295350)
AUTHORS Director-General, Biotechnology Center, Aoki, K., Oguchi, A., Nagai, Y., Asano, K., Iwama, N., Baba, T., Kuroda, M., Hiramatsu, K. and Kikuchi, H.
TITLE Direct Submission
JOURNAL Submitted (06-MAR-2002) Director-General, Biotechnology Center, National Institute of Technology and Evaluation, Biotechnology Center; 2Chome 49-10 Nishihara, Shibuya-ku, Tokyo 151-0066, Japan (E-mail:bioelite.go.jp, URL: <http://www.bio.nite.go.jp/>, Tel: 81-3-3481-1933, Fax: 81-3-3481-8424)
FEATURES
 source
 1. 295350
 /organism="Staphylococcus aureus subsp. aureus MW2"
 /mol_type="genomic DNA"
 /strain="MW2"
 /db_xref="taxon:196620"
 301..1254
 /gene="MW1096"
 301..1254
 /gene="MW1096"
 /note="ORFID:MW1096"
 /codon_start=1
 /transl_table=11
 /product="hypothetical protein"
 /protein_id="BAB94961.1"
 /db_xref="GI:21204264"
 /translation="MKKTLCGLLIMLVAGSPGNHKLSSKSESKQETVKES
 EEKDPLEKYEIEIKMGKIDKAPSLDKLDPLMTKSFNSKGIQGWKDYKELMGKV
 ELADYRTKDSGSSIKDVAFFKGGKIKRVIETHDDVKQDYVYDPDGKKIGNS
 NTFVFAEIMTKYKDGKLVYASVEPGSVYIHKDDAIKYDDYSKLKLSQLTKLDHFKP
 VYPSVAQIKSEGVPLTSVSTHSGDKTDEVLPALEYFTFSKPNYEDKSNPDPKVLN
 LVHMDFLNASDQNAHFVYLSKYIKESYENYETASDSSLK"
 complement(1417..1695)
 /gene="MW1097"
 /note="MW1097"
 /gene="MW1097"
 /note="ORFID:MW1097"
 /codon_start=1
 /transl_table=11
 /product="conserved hypothetical protein"
 /protein_id="BAB94962.1"
 /db_xref="GI:21204265"
 /translation="MPELVGILIKREANVLSNSTGKEMQVKNVYVILLALFLGSFG
 IHPYAGKNMQGILHLIFCWGIPHLAISAVITVPKPADEQGNVL"
 1919..2407
 /gene="MW1098"
 1919..2407
 /gene="MW1098"
 /note="ORFID:MW1098"
 hypothetical protein, similar to polypeptide deformylase"
 /codon_start=1
 /transl_table=11
 /protein_id="BAB94963.1"
 /db_xref="GI:21204266"
 /translation="MAIKLVPAASHPLTKKAQAVIKFDDSLKRLQLDLEDTWYAOEA
 AGLCAQINSQVLAIDMEMEGLQLVNPVKIISQSNETITDLEGSITLDPVYGEVTR"

SKMIYVESYDVNGKNKVELTAHEDVARMILHIIDQMNGIPFTTERADRILTDKEVEAYFI
 ND"
 2400..3335
 /gene="MW1099"
 2400..3335
 /gene="MW1099"
 /note="ORFID:MW1099"
 /codon_start=1
 /transl_table=11
 /product="methionyl-tRNA formyltransferase"
 /protein_id="BAB94964.1"
 /db_xref="GI:21204267"
 /translation="MTKIFMGTPDFSTVLEMLIAEHDVIAVTPQDPVGRKRVMT
 PPVKVAMKYDLVYQPEKLSGSELEQLQLDVLIVTAAGQLLPESLIALPKLG
 AINVASLLPKYRGAPLHQALIDGEGTGTIMYVKKLDAGNIISQOAIKLENDN
 VGMHDKLSVLGADLKTLEFSLIEGTNESVPDDTQATFASNIRREDERINWKPGR
 QVFNQIRGLSPWPVAYTTMDPTNLKIYDAELVETNKINEPGTITTIETTKAILIVATNDN
 EAVAIDKMLQLAGKRMALANYLSGAQNTLVGKKLI"
 3332..4639
 /gene="MW1100"
 3332..4639
 /gene="MW1100"
 /note="ORFID:MW1100"
 hypothetical protein, similar to RNA-binding Sun protein"
 /codon_start=1
 /transl_table=11
 /protein_id="BAB94965.1"
 /db_xref="GI:21204268"
 /translation="MIENVRSLAFDTIQDILNEGAYSNLRINEVLSENELNMDKALF
 TEIVYGVTKRKFTLDFYLPKPFVKTIKAWVRQLLWMSIYQYVYLDKVPNHAINAEVE
 TAKERGYNHGVNNGILRTMMRSDLPDFNEIADPKRMAIEYSPMKWIIDHWATHYG
 LEKTTILQSELETSTTVRANLTRASLDDIIEKLQDEGYDVEKDHDLPYCLHIGSQP
 IHSRSFKDSFVSIQKSMFVAHIMVDRHDHVLVDACSAPGKAKACHAEVILMPGQV
 DASDIHDKIDILNINIKKRLTNKAFQHDATKPYDSTYDKILVDAPCSGUGVWRHK
 PEIKYTSQKHIESLVEQLLELVNKNVKGIGIYKSTCTIEQLENEENVITFLKN
 NKNFEPFQHPITGELVKTQLMPQDFNSDGFITTKIKRKN"
 4642..5736
 /gene="MW1101"
 4642..5736
 /gene="MW1101"
 /note="ORFID:MW1101"
 /codon_start=1
 /transl_table=11
 /product="conserved hypothetical protein"
 /protein_id="BAB94966.1"
 /db_xref="GI:21204269"
 /translation="MITAEKKKKFLPNFKQSIYSLSRFDDEMKNLVBQGOQKFRK
 QIFEWLYOKRVDSDIDMTNLSKDLQLKDNFTVTLTTVVKQESKDGITIKELYEQD
 GYTIETVLMRHDYGNVCVTVQVGRIGCTFCASTLGLKRNLEAGEIVSOVLTVOKA
 LDATERVSIQVIMGIGEPFENYDEMDFLRIVNDNSINIGARHITVSTSGILPRIY
 DPAEDIQINFVSLHAADKVRSLMPINRANVNEKLEAIQYQOEKTRNRVTFRYG
 LFGYNDQLEHARELAHLIKGLNCHVNLIPVNVHPERNYVTKAKNDIFKPEKELKRLG
 INATIRREGQSDIDAACQGLRAKQRVETR"
 5743..6486
 /gene="MW1102"
 5743..6486
 /gene="MW1102"
 /note="ORFID:MW1102"
 /codon_start=1
 /transl_table=11
 /product="conserved hypothetical protein"
 /protein_id="BAB94967.1"
 /db_xref="GI:21204270"
 /translation="MLEAQFTDTQHRDKNEDAGGIFYNRTNQQLLVLCDSGGGHKA
 GEVASKFVTDLKSFRFPAENLIEEHAENLRNNKDNFNFOLYHAQENAEYKMGMTT
 CVCALVFEKSVIVANVDSRAVINSRQEQITSDHSFVNHLVLTQITPEAFTHPQ
 ENLITKVMGTDKRVPDLFIKRLNFVYLLNSDGLTDVVKDNEIKRLLVKEGTIEDH
 GQQLQALDNDHNSKDNVTFILAAIEGDKV"
 6483..8477
 /gene="MW1103"
 6483..8477
 /gene="MW1103"
 /note="ORFID:MW1103"

hypothetical protein, similar to protein kinase"

```
/codon_start=1
/transl_table=1
/protein_id="BAB94968.1"
/db_xref="GI:21204271"
/translation="MWIGIIINERYKIVDKLGGGCMSTVYLAEDTILNLIKVAIKAFIP
PREKETUKRERERVHNSQLSHQNIIVSMIDVDEDDCYILVMEYIEGPTLSEVIESH
GPLSDVTAINTFQNLGDIKHAHDMRIVHRDIKQNIILIDNSNKLKIPDFGIAKALSE
TSLTQTNHVLQVQVPEQAKGEATDECTDIYSIGIVLYEMLVGEPPFNGETAVSIA
IKHTQSDVPNTVDVRKDIPOSLSNVILRATKOKANRYKTIQEMKDDLSVLHENRA
NEVYELDKMTIAVPLAKEDLAKHISEHKSNOPTQVPIVNGPAHQFOFKPEG
TVPEPKKSKTRKIVLKSLLFSLMLIALVSPVAMAMGKXYEETPPDVIGKSVKEAQ
IPKNKLKLGKISRSYDKYPENELIKTTPNGERKVRGSDVVDVVISGPKVKMPNV
IGLPEALQKLKSLGLKDVITIEKYNNOAPKGYIANOSVTAIEAIDHSNLIKLYES
LGIHQVYVEDFEHKSFSKAKALEKGFVSKESYDDIDEGDVVISOSPKGKSVDEG
STISFVSGKKSOSSDVKTTTESVDVPTGKNDKSKQVKVYIKDKNDGSTEKGSFD
ITSQDRIPIRIEKGKTASYIVKVDGKTVAEKSVDV"
```

gene

8705..9580

CDS

8705..9580

```
/gene="MW1104"
/translation="ORFID:MW1104"
/codon_start=1
/transl_table=1
/product="conserved hypothetical protein"
/protein_id="BAB94969.1"
/db_xref="GI:21204272"
/translation="MKTGRIVKSIISGVYQVDWNGERFNTKPRGLFRKKKFFSVVGDIV
EFVQNIINEGYIHOVFERKNEIKPPVSNIDTIVMSAVEPNSTOLLDRFLVIAHS
YQANARVLTKKDKTPISEKQFNEILLKIYENIGYETEFIGNDDRRKKIVENAWFAGLI
VLSGSGVGFSLNHYPELNLNDISKLNKRGKHTTRHVELFQRNGYIADTPGF
SALDPGDHDKDEIKYFLELARYGETCKFRNCNHIKEPNCNVKHQLEIGNIAQFRYDH
YLOLFNEISNRKRVY"
9581..10225
/gene="cfxE"
9581..10225
/gene="cfxE"
/translation="ORFID:MW1105
ribulose-5-phosphate 3-epimerase homolog"
/codon_start=1
/transl_table=1
/protein_id="BAB94970.1"
/db_xref="GI:21204273"
/translation="MTKLPSLLSVDFLDLQHELRLEAEAGVGVHFDVMDGQVPNI
SIGPLIDAVKRGITLPIIDVHLMENPEKTIASPAEHGADMISIHVESTPHIRAIOM
IKHLDKKAGVINPTGPTISOIEPILDIVDVLVMTVNPGRGSGFIDQCVEKIAAGLA
IKMERQLNFDIEVDGVTNTAKVCVENGATMLVTGSPFFKQEDYKKVTOALKG"
10232..10873
/gene="MW1106"
10232..10873
/gene="MW1106"
/translation="ORFID:MW1106"
/codon_start=1
```

gene

10232..10873

CDS

10232..10873

Query Match 98.3%; Score 723.2; DB 1; Length 295350;
Best Local Similarity 98.9%; Pred. No. 6.5e-123;
Matches 728; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

```
QY 1 ATTTACGTTTGTCTATCATACCAACGCCAGCAACTGTTTCAATCAATTTGTC 60
DB ATTTACGTTTGTCTATCATACCAACGCCAGCAACTGTTTCAATCAATTTGTC 230642
QY 61 ACGATCGTATGTTTAACTGTTTAAATATCTTATATAAATCATCTACGACATTTGT 120
DB ACGATCGTATGTTTAACTGTTTAAATATCTTATATAAATCATCTACGACATTTGT 230702
QY 121 TTCTACTTCACTATTATACCCCAATCATGATTTTAAATTTGTTCCCGTTGCATAACATG 180
DB TTCTACTTCACTATTATACCCCAATCATGATTTTAAATTTGTTCCCGTTGCATAACATG 230762
QY 181 GTTTTATTTTTCAGTAGAATATAGTAATATCATCTCTGTTTGTGTAATTTCAATTC 240
DB GTTTTATTTTTCAGTAGAATATAGTAATATCATCTCTGTTTGTGTAATTTCAATTC 230822
```



```
source
1. .303750
/organism="Staphylococcus aureus subsp. aureus N315"
/moi_type="genomic DNA"
/strain="N315"
/sub_species="aureus N315"
/db_xref="taxon:158879"
202. .1155
/gene="SA1056"
CDS
202. .1155
/gene="SA1056"
/notes="ORFID:SA1056"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAB42308.1"
/db_xref="GI:13701013"
/translations="MKRTKGLLLIMLVVAGSGFGNHLKSSKSEESKOETVKKES
EEKDPLEKYEIEIKMGKIGIDAPSLDKLDPLMTKSGFTNSKGIOGWKYELMGKV
ELADYRPTKSGSSIKVDAPFKGKGKIRKVIETHDDVKQDVYVYVDPDGKKGNS
NTPVFAEIMTKYKDGKLYVASVEPGSYVHKDDAJKYDDYSKLKLSQTLKLDHPKP
VPYSAQIKSGFGLPLTSVSWTHGSKDTEVLDPALAYFTFSPKNTYEDKSNPDPKVLN
LVHMDLFIASSDFGNAHFVVLISKYIEKYESNYETASDSDLK"
/gene="SA1057"
complement(1318. .1521)
complement(1318. .1521)
/gene="SA1057"
CDS
/gene="SA1057"
/notes="ORFID:SA1057"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAB42309.1"
/db_xref="GI:13701014"
/translations="MQNKKVIYIILLFLGSGFIHFKYAGKNQGIHLHIFCWGTGIPH
ILAIIASAVITVFKPADEQGNVTL"
1820. .2308
/gene="SA1058"
CDS
1820. .2308
/gene="SA1058"
/notes="ORFID:SA1058"
hypothetical protein, similar to polypeptide deformylase"
/codon_start=1
/transl_table=11
/protein_id="BAB42310.1"
/db_xref="GI:13701015"
/translations="MAIKKLVPASHILTKQAQVTKPDDSLKRLLOLQLEDNTWAOEA
AGLCAPQINQSLQVAIIDMEGELLQVNPKIISQSNETITDLEGSITILPVDVGEVTR
SKMIVVESYDVNGKNVELTAHEDVARMILHIIDQNGIIPFTERADRILTDKEVEAYFI
ND"
2301. .3236
/gene="SA1059"
CDS
2301. .3236
/gene="SA1059"
/notes="ORFID:SA1059"
/codon_start=1
/transl_table=11
/product="methionyl-tRNA formyltransferase"
/protein_id="BAB42311.1"
/db_xref="GI:13701016"
/translations="MTKIIIPMGTFPDTTVMLELIAEHVDVIAVVTQDPVGRKRVMT
PPVKKVAMKYDLPVYQPEKSGSELEQLQDLDVLIIVTAAFGQLLPESLLALPKLG
AINVHASLLPKYRGGAPIHOAIIIDGQETGITIMYVKKLDAGNIIISQAIKTEENDN
VGTMDKLSVLGADLLKETLPSIIIEGTNESVPQDDTQATFASNIRREDRISWNKPCR
QVFNIRGLSPFWPVAYTMDTDLNLIYDAELVETKNINBPGTIIETTKKAILVAITNDN
EAVAIDMQLAGKMLAANYLSGAQNTLVGKKLI"
3155. .4540
/gene="SA1060"
CDS
3155. .4540
/gene="SA1060"
/notes="ORFID:SA1060"
hypothetical protein, similar to RNA-binding Sun protein"
/codon_start=1
/transl_table=11
/protein_id="BAB42312.1"

/db_xref="GI:13701017"
/translations="MSMEKENVSCOLFVKCAKHTSRETYMLENVRSIAPDTIODILN
EGAYNLKRINEVSENLNAMDALFTIIVYGTVRKRYTLDYLPKFPVKTKIKAWVRQ
LLWMSIYQYLDKPNHAINAEVAIEAKERGYYHNGVNGILRTMMKSDLPDFNEI
ADPKRMAIEYMPKWIIDHWATHYGLBEETILOSFLIETTSITVRANITRASLDDII
EKLDEGYDVEKDHLPYCLHITGGQPIIHSRSFKDGFVSIDQKSMFVAHIKMWDRHD
HVLDCASPGGKACHAEVLMPGQVDSDIHDHKIDILINFINIKLRLTNIKAFORHDA
TKPYDKYDKILVADAPCSGLGVMRKPEIKYTKSQKHIESLVELQLEILLENKKNVKI
GGELIYSTCTIELENENVIYFLKKNKRFEPFQHPITGSLVTKLQIMPQDFNSDG
FFITIKKRDN"
4543. .5637
/gene="SA1061"
CDS
4543. .5637
/gene="SA1061"
/notes="ORFID:SA1061"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAB42313.1"
/db_xref="GI:13701018"
/translations="MITAEKKKNKFLPNFDKQSIYSILRPFDEMQNWLVEQOQKFRK
QIFEWLYQKRVDYSIDEMTNLSKDLQLLNKDNFTVTTLTVVVKQESKDGITKFLFELQD
GYTIIETLMRHDYSGNSVCTVQVGRIGCTFCASTLGLKRLNLEAGEIVSQLTVKQA
LDATERVSOIVIMGIDYSGIFEPFENYDEMDFLRIVNDNLSNIGRAHITVTSGITPRIY
DFADEDIQINFAVSLHAAKDEVRSLMPINRAYNVEKLEAIQYYQEKTRRVTPYEG
LEGVNDQLEHARELAHLIKGLCHVNLI PVNHVPERNVYVTKANDIIFKEKELKRLG
INATIRREGGSDIDAACGQRAKERQVETR"
5644. .6387
/gene="SA1062"
CDS
5644. .6387
/gene="SA1062"
/notes="ORFID:SA1062"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAB42314.1"
/db_xref="GI:13701019"
/translations="MLEAQFFDTTQCHRDKNEDAGGIFYNQTNQQLLVLCDCMGHKA
GVSASFVTEDELKSRPEAENLIEHQAENLNRNINKDINFQIHYHAQENAEYKMGTT
GVCALVFKEKSVIANVGSRAVINSRQIEITSDHSFVNHLVLTGQITPBEAFTHPQ
RNIIITKVMGTDKRVSPLFIKRLNFYDILLNSDGLTDYVKDNEIKRLLVREGTTIEDH
GDQLMQLADNHSKDNVTFILAAIEGDKV"
6384. .8378
/gene="SA1063"
CDS
6384. .8378
/gene="SA1063"
/notes="ORFID:SA1063"
/codon_start=1
/transl_table=11
/product="protein kinase"
/protein_id="BAB42315.1"
/db_xref="GI:13701020"
/translations="MIGKIINERYKIVDKLGGGMSVYLAEDTILNIKVAIKAFIP
PREKEETLKRPEREVHNSQLSHONIVSMIDVEEDDCYLVMEYIEGPTLSEYIESH
GPLSVDTAINTNQILDGIKHADMRIVHRDIKPNQILIDSNKTLKIPDFGAKALSE
TSLTQTNHVLGVTVQVFSPEQAKGATDECTDIYSIGIVLYEMLVGPFPNGETAWSIA
IKHIQDSVNVITDVKRIQPSLNSVILRATKEDKANRYKTIQEMKDDLSLVSHENRA
NEDVELDKPDKTIAVPLKEDDLAKHISEHKNQPKRETTQVPIVNGFAHQFQKPEG
TVYEPKPKKSTRIKVLISLI FSLMIALVSVFAMFGNKYEETPDVIGSKYKAEAQ
IGLPKKEALQKLKSLGDKVTTEKVVNQAQPKYIQAQSVTANTEIAIHDSNIKLYES
LGIQVYVDFEHKFSKAKALEKGPVKESKEEYSDIDSGDVIQSQPKGSVDGEG
STISFVSGKKSDDSDVKTTTSDVDPYTGKNDKRSQKRVYIKONDNQDGTSEKSGFD
ITSDQRIDIPRIEIKGTASYIVKVDGKTVAEKEVSYYDDI"
8606. .9481
/gene="SA1064"
CDS
8606. .9481
/gene="SA1064"
/notes="ORFID:SA1064"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAB42316.1"
```

```
/db_xref="GI:13701021"
/translation="MKTRIVKSIQVGVQVNGERFNTKPRGLFRKKKSPVVGDIV
EFVQNIQEGYIHQVFERKELKPPVSNIDTLVMSAVEPNFSQLLDRFLVIAHS
YQLNARVLTKKDKTPEKQFIELNELLKIYENIGVETFGINDDRKKIYEAWPAGLI
VLSGSGVGKSTFLNHYPELNLNLTNDISKSLNRGKHTRHVELFERONGYIADTQGF
SALDPDHDKDEIKDYFLELNRYSCTCKFRNCNHIKEPNCNVKQLEIGNIAQFRYDH
YLOLFNELSNKRVY"
9482..10126
/gene="cfxE"
9482..10126
/gene="cfxE"
/feature="ORFID:SA1065
ribulose-5-phosphate 3-epimerase homolog"
/codon_start=1
/transl_table=11
/protein_id="BAB42317.1"
/db_xref="GI:13701022"
/translation="MTKLYPSLLSVDPLDLOHELKRLBEGVDGVHFDVMDQFVNT
SIGLPILDVAVKGTPLPIDVHLMENPEKYIASFAEHGADMISHVSTPCHIRAIOM
IKHLDKAGVINGTPIISQLEPILDIYVYLVMTVPNGFGQSFIDOCVEKIALNA
IKMERQLNFDIEVDGVTNDTAKVCVENGATMLVTGSPFKQEDYKKVTQALKG"
Query Match 94.3%; Score 694; DB 1; Length 303750;
Best Local Similarity 97.8%; Pred. No. 1.4e-117;
Matches 720; Conservative 0; Mismatches 0; Indels 16; Gaps 1;
Qy 1 ATTTACGTTTTGTTCATCGTATCACATACCCACGCGACGAACTGTTTCAATCATTTTGTGTC 60
Db 227532 ATTTACGTTTTGTTCATCGTATCACATACCCACGCGACGAACTGTTTCAATCATTTTGTGTC 227591
Qy 61 ACGATCGTATGGTTTTAACTGTTTTCGTAATAATCTTATATAAATCTACGACATTTGT 120
Db 227592 ACGATCGTATGGTTTTAACTGTTTTCGTAATAATCTTATATAAATCTACGACATTTGT 227651
Qy 121 TTCTACTTCACTATTATACCCACATACATGATTTTAAATTTGTTCCCGTTCGATAACATG 180
Db 227652 TTCTACTTCACTATTATACCCACATACATGATTTTAAATTTGTTCCCGTTCGATAACATG 227711
Qy 181 GTTTTATTTTTCAGCTAGAGATATAGTAAATCATACACTCTGTTTGTGTAATTCATTTTC 240
Db 227712 GTTTTATTTTTCAGCTAGAGATATAGTAAATCATACACTCTGTTTGTGTAATTCATTTTC 227771
Qy 241 TCGCGCATTTACCGTCACCTTTAAAGCGTTCTTATCAATGTAATACCGTTGACATCGAT 300
Db 227772 TCGCGCATTTACCGTCACCTTTAAAGCGTT-----CGTTGACATCGAT 227815
Qy 301 AATATCCTTTTGTGCTGACGACGTAAATTTGACGAAATTTCTGCTAAAGTTCTTCAAT 360
Db 227816 AATATCCTTTTGTGCTGACGACGTAAATTTGACGAAATTTCTGCTAAAGTTCTTCAAT 227875
Qy 361 ATCAACCGGCTTAACATATATCGTCTGACCGTAATCAAGCCGACGAACTTTGTGTCATA 420
Db 227876 ATCAACCGGCTTAACATATATCGTCTGACCGTAATCAAGCCGACGAACTTTGTGTCATA 227935
Qy 421 CGTATCACTTTTTCGCTGTAATTAATGATAGGTGTAGATTTGTTGTTCTAAATTTGCG 480
Db 227936 CGTATCACTTTTTCGCTGTAATTAATGATAGGTGTAGATTTGTTGTTCTAAATTTGCG 227995
Qy 481 ACAATTTCTAAGCCATTAAATGACGCAACATTAATTAATTAATTAATTAATTAATTAAT 540
Db 227996 ACAATTTCTAAGCCATTAAATGACGCAACATTAATTAATTAATTAATTAATTAATTAAT 228055
Qy 541 ATGGCTAAGCGCTTTATCAACCGCTTTGTCATCATCATCTGTCGTCATCATTTGTAAT 600
Db 228056 ATGGCTAAGCGCTTTATCAACCGCTTTGTCATCATCATCTGTCGTCATCATTTGTAAT 228115
Qy 601 TTCATGTGTGAGTCCAAATTCAGAAATCTTGTGTAAGTTTGTTCATCTTCTACTATTA 660
Db 228116 TTCATGTGTGAGTCCAAATTCAGAAATCTTGTGTAAGTTTGTTCATCTTCTACTATTA 228175
Qy 661 AATTTGGCTCATTTGTACACCTCATATACACCTTTTCTAATTAAGGTAATATATTTAA 720
Db 228176 AATTTGGCTCATTTGTACACCTCATATATACACCTTTTCTAATTAAGGTAATATATTTAA 228235
```

```
Qy 721 ATTTTGTATTTTCTA 736
Db 228236 ATTTTGTATTTTCTA 228251

RESULT 11
AX621578/c 657 bp DNA linear PAT 20-FEB-2003
LOCUS AX621578 Sequence 4541 from Patent WO02094868.
DEFINITION AX621578
ACCESSION AX621578
VERSION AX621578.1 GI:28450108
KEYWORDS Staphylococcus aureus
SOURCE Staphylococcus aureus
ORGANISM Staphylococcus aureus
REFERENCE 1
AUTHORS Masignani, V.C., Mora, M.C. and Scarselli, M.C.
TITLE Staphylococcus aureus proteins and nucleic acids
JOURNAL Patent: WO 02094868-A 4541 28-NOV-2002;
Chiron Spa (IT)
FEATURES
Location/Qualifiers
source 1..657
/organism="Staphylococcus aureus"
/mol_type="unassigned DNA"
/db_xref="taxon:1280"

ORIGIN
Query Match 89.3%; Score 657; DB 6; Length 657;
Best Local Similarity 100.0%; Pred. No. 2.5e-110;
Matches 657; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 16 TCGTATCACATACCCAAACGCGACGAACTGTTTCAATCATTTTGTACGATCGTATGTTT 75
Db 657 TCGTATCACATACCCAAACGCGACGAACTGTTTCAATCATTTTGTACGATCGTATGTTT 598
Qy 76 TAACCTGTTTCGTAATAATCTTATATAAATCTTACGACATTTGTTTCTACTTCACTATT 135
Db 597 TAACCTGTTTCGTAATAATCTTATATAAATCTTACGACATTTGTTTCTACTTCACTATT 538
Qy 136 ATACCCCATACATGATTTAAATTTGTTCCCGTTCGATACATGTTTATTTTTCAGC 195
Db 537 ATACCCCATACATGATTTAAATTTGTTCCCGTTCGATACATGTTTATTTTTCAGC 478
Qy 196 TAGAAGATATAGTAAATCATCTCTGTTTGTGTTTAAATTTCTGCGCATTTACCGT 255
Db 477 TAGAAGATATAGTAAATCATCTCTGTTTGTGTTTAAATTTCTGCGCATTTACCGT 418
Qy 256 CACTTTAAAGCGTTCTTATCAATTTGTAATACCGTTGACATCGATATATCCTTTTGTGG 315
Db 417 CACTTTAAAGCGTTCTTATCAATTTGTAATACCGTTGACATCGATATATCCTTTTGTGG 358
Qy 316 CTGACGACGTAATAATTCGACGAAATTTCTGCTAAAGTTCTTCAATATCAACGCGTTAAC 375
Db 357 CTGACGACGTAATAATTCGACGAAATTTCTGCTAAAGTTCTTCAATATCAACGCGTTAAC 298
Qy 376 TATATAATCTGCTGACCGTAATCAAGCCGACGAACTTTGTGTCATCGATGTTTTCGC 435
Db 297 TATATAATCTGCTGACCGTAATCAAGCCGACGAACTTTGTGTCATCGATGTTTTCGC 238
Qy 436 TGTAAATTTAATGATAGGTGTAGATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 495
Db 237 TGTAAATTTAATGATAGGTGTAGATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 178
Qy 496 ATTAATTGACGGCAACATTAATAATCTAATATGATTAATCATAGTAATGCGTAAAGCGTTT 555
Db 177 ATTAATTGACGGCAACATTAATAATCTAATATGATTAATCATAGTAATGCGTAAAGCGTTT 118
Qy 556 ATCTAAACCGCTTGTGTCATCATCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 615
Db 117 ATCTAAACCGCTTGTGTCATCATCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 58
Qy 616 CAATTCAAGAAATCTTGTCTAAGTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 672
```


Qy	122	TCTACTTCACTATTATAACCCCATACATGATTTAAAAATTTGTTCCGGTTGCATAACATCGG	181
Db	208780	TCTACTTCACTATTATAACCCCATACGATCGAGAAATTTGTTACGTTTGCATGACGTGG	208839
Qy	182	TTTTTATTTTTCAGCTAGAGATATAGTAATATCATCTCTGTTTTTGTAAATTCATTTCT	241
Db	208840	TTACGATTTTTCAGCTAGAGATATAGTAATATCATCTCTGTTTTTGTAAATTCATTTCT	208899
Qy	242	GCGCCATTTTACCGTCACTTTTAAAGCGTTCTTATCAATTTGTAATACCGTTGACATCGATA	301
Db	208900	TGGCCATTAAACGTAACCTTTTAAAGGCACTCTTTATCAATGATATACCATGATATCTAAA	208959
Qy	302	ATATCTTTTGTGGCTGACGACGTAATAATTCGACGAATTTCTGTCTAAAGTCTTCTCAATA	361
Db	208960	ACATCTTTTATCTGGCTGCTGCGCAATACCGCTCTTATCTTTTTCGCAAGCAATTTCTCTATA	209019
Qy	362	TCAAACGGCTTAACATATATAATCGTCTGACCGTAATCAAGCCACCAACCTTTTTCATAC	421
Db	209020	TCAAAGGGTTTAAATGATGATCATCTGCCCATAGTCCCAATCCAGCTACTTTTATCATAT	209079
Qy	422	GTATCACTTTTTCGCTCTGTAATTAATGATAGGTAGATTTGTTGTTCTTAATTTTTCGGA	481
Db	209080	GTCTCGCTTTTTCGAGTAATGATGATTAATTTGGAGTAGTTGTTTTTTCGCAATTTGTTCTA	209139
Qy	482	CAAAATTTTAAAGCATTAAATGACGCAACATTAAATCTTAATATGATTAATATAGTAA	541
Db	209140	CAAAATTTTACGACATTTTATTTTGAAGCATTAAAGTCTTAATATATATAAATCATAGGCG	209199
Qy	542	TGCTTAAGCGCTTTATCTAAACCGTCTTGTCCATCATCTCTGTCCACATTTCTAATTT	601
Db	209200	TTAGATAATGCTTTTATCCAAACCCACCTTCCATCATCTCTTATATCAACAGATTAATTT	209259
Qy	602	TCATGTGTGAGTTTCAATTCAGAAATCTTGTCAAGTTTGTTCATCTTCTTCTACTATAA	661
Db	209260	TCATGAGTTAACTCAAGCTCTATAATCTAGCAGATTTTGTCTCATCTTCTTCAATTA	209319
Qy	662	ATTTGCGTCAATTTGTACACCTCATATTACGACTTTTTTCTAATAA	705
Db	209320	ATATTTGTCATATTTCACCTCAGCTACATCTTAAACAATTA	209363
RESULT 13			
LOCUS	AF269746	3099 bp DNA linear	BCT 01-AUG-2000
DEFINITION	Staphylococcus epidermidis strain SRI clone step.1021e09 genomic sequence.		
ACCESSION	AF269746		
VERSION	AF269746.1 GI:9623642		
KEYWORDS	Staphylococcus epidermidis		
SOURCE	Staphylococcus epidermidis		
ORGANISM	Bacteria; Firmicutes; Bacillales; Staphylococcus.		
REFERENCE	1. (bases 1 to 3099)		
AUTHORS	Kimberly, W.J., Taylor, J. David., Nelsen, A.J., Godlevski, M.M., Rubino, M.A., Nelsen, F.J., Rivers, P.R., Torruella-Miller, I., Liskenbee, S., Ahsanti, C., Altschuler, G., Mamo, L., Shepherd, N.S., Fuchs, R., Fleming, T., Guan, X., Du, L., Cain, D.H., Miller, G.S. and Furdon, P.J.		
TITLE	Transposon-mediated sequencing of the Staphylococcus epidermidis genome		
JOURNAL	Unpublished		
REFERENCE	2. (bases 1 to 3099)		
AUTHORS	Taylor, J. David., Kimberley, W.J., Nelsen, A.J., Godlevski, M.M., Rubino, M.A., Nelsen, F.J., Rivers, P.R., Torruella-Miller, I., Liskenbee, S., Ahsanti, C., Altschuler, G., Mamo, L., Shepherd, N.S., Fuchs, R., Fleming, T., Guan, X., Du, L., Cain, D.H., Miller, G.S. and Furdon, P.J.		
TITLE	Direct Submission		
JOURNAL	Submitted (22-MAY-2000) Departments of Genomic Sciences and Bioinformatics, Genetics Directorate, Glaxo Wellcome, Inc., 5 Moore Drive, Research Triangle Park, North Carolina 27709-3398, USA		
FEATURES	Location/Qualifiers		
source	1. .3099		

gene	/transl_table=11			
	/product="rimM protein"			
	/protein_id="AA004511.1"			
CDS	/db_xref="GI:27315377"			
	/translation="MEVEVGQIVNTHGIGKEGVKXNSDFTETRFQPCQEQLLVKNHNT EIVYTVASRIHKGPHMLPEGINNINEIHLKGDYIYQERDHDQIELGEHYYYSDI IGCTVFDODDTPIGRVINIFETGANDVWVVGKEKEYLPIYIEDVVVKDIDIEKTIKIT PMEILLD"			
	9186. .9923			
	/locus_tag="SE0915"			
CDS	9186. .9923			
	/locus_tag="SE0915"			
	/codon_start=1			
	/transl_table=11			
	/product="trRNA (guanine-N1) -methyltransferase"			
	/protein_id="AA004512.1"			
	/db_xref="GI:27315378"			
	/translation="MKIDYLTLPFPMFEGVLNHSILKRAQDKGIINVTINPRDYSIN KHNQVDYFPGGQGVHLKPEVPFNAMEDINNEHTRVLDMCPQGRPTQBIAGELSE AKHIVDICHGEYEDIRIKHLVDTBEISMGDYLTCGELPAMTMDAIVRLIPVLGN QASHQDQSFSDGLLEFPQTRPREYKNMSVPEVLISGNHAHIDOWRHEQKLIRTYEKR PDLEQVPLTEKOREILETYKKKLKND"			
	10029. .10379			
gene	/locus_tag="SE0916"			
CDS	10029. .10379			
	/locus_tag="SE0916"			
	/codon_start=1			
	/transl_table=11			
	/product="50S ribosomal protein L19"			
	/protein_id="AA004513.1"			
	/db_xref="GI:27315379"			
	/translation="MSNHKLIEAVTKSKQLRDLTPFRSGDRLRVHVRIVEGSRRIQV FGVTKRRGGISSETFTVKRISSGVGVERTPLTPKIEKIEVRKRGKVRRAKLYL RSLRGKAARIQER"			
	complement (10910. .13513)			
gene	/locus_tag="SE0917"			
CDS	complement (10910. .13513)			
	/locus_tag="SE0917"			
	/codon_start=1			
	/transl_table=11			
	/product="conserved hypothetical protein"			
	/protein_id="AA004514.1"			
	/db_xref="GI:27315380"			
	/translation="MKINKNPKIKRFTLLVLVSCILSPILYGLPLYQFIKGVFSG SGDFQPMQFOMHLYEHFTSLKGFDSAGLDGVKSLVYVYSLSLMWNFMSMW ILQETNVNPHDLSFTWINQIMAYVUTVITFESFLVSLRKPAPKFWATITLYGM STVVTYNTWSPYGNLLMLPMSLWAIERPFKERRKIGWFIAPATLFSNPFYSYE AVIVGFYIRFAIPEKDIIVARQGYILVCATLLSLYGLYQVGVSSFLDNRIL QNPKITFTFTNFYINIFADGVYITSTIAILFCFKLYQHYKKJLFAIATWIL IGVFIALLSRSVKMVFALLILIVLRFIKYKSLTRTSMVLLVFLFAAQVLT IYHNQTSYSSIFDGDILKYIDQTLQINMPVDKNSYRILNKLMSLMDVQDLR HPDLNMGYFKKKELITDKDQIHSVNTINYPASHITNKIYDARKLKSLDREQAM LKGVNLHKSQANTDPKPNLLSNKQNLNHNAMIDSKHLVKQHGNGVTLNLPRI VKNYKDMYIEMQNDVLLSPKSKHGVNVEYSQERNLSYKRYRFSVPTMRKASNLN IAKMSGVYRQKVIYGVNYQTLKASQOLQPVKKXERNGFTIHKKKHGHYLVLPV VAKGGMHANGKPLKQVQNGIMTIPVKESQTKILSLYPPFPYLLITVSCIGIIL SLFTHYVRK"			
	complement (13506. .15779)			
gene				
Query Match	56.1%; Score 412.8; DB 1; Length 300892;			
Best Local Similarity	74.1%; Pred. No. 3e-66;			
Matches 522; Conservative	0; Mismatches 182; Indels 0; Gaps 0;			
2	TTTACGTTTGTGTCATCGATACATACCAACGCGACCAATCTTCAATCATTTTGTCA 61			
208660	TTGGCGCTTAATCATCGAATACATACCTACGCGACCTACTGTTCTATGATTTTCT 208719			
62	CGATCGTATGTTTAACTGTTTTCGTAATATCTTATATAACATCTACGACATTTGTT 121			
208720	TTATTAAGAGTTTGTAGTTTATAGTAATAACGAATATAACATCAACGACATTCGTT 208779			

/transl_table=11
/product="rimM protein"
/protein_id="AA004511.1"
/db_xref="GI:27315377"
/translation="MEVEVGQIVNTHGKGVKVSNDFTETFPQGEQLVKNHT
EIVTVAYRHKHGFMDYFQERDQDQIEGEHEYYSDI
IGCTVFDQDDPIGRVINIFETGADVNVVKEEYLIPIYEDVVKDIDIBETIKIT
PMEGLD"
9186..9923
/locus_tag="SE0915"
9186..9923
/locus_tag="SE0915"
/codon_start=1
/transl_table=11
/product="crRNA (guanine-N1) -methyltransferase"
/protein_id="AA004512.1"
/db_xref="GI:27315378"
/translation="MKDIDYLPFEMFEGVLNHSILKRAQDKGIINVTINFRYSIN
KHNQDDIPFGGGQVMDLKPFPVNAMEIDNNEHRTVILMCPQGRPTQBIAGELSE
AKHIVICFHYEGYDRIKRLVTDETSMGDYLTGGELPAMTMDAIVRLIPGLGN
QASHQDDSFSDGLLEFPQYTRPREYKNMSVPEVLLSGNHAHIDQWRHEQKLIRTYER
PDLLEQVPLTEKREILLETYYKKLKD"
10029..10379
/locus_tag="SE0916"
10029..10379
/locus_tag="SE0916"
/codon_start=1
/transl_table=11
/product="50S ribosomal protein L19"
/protein_id="AA004513.1"
/db_xref="GI:27315379"
/translation="MSNHLJEAATKSQLRDLTFRSGDTRVHRVIRVEGSRRIQV
PEGVVKRGGIGSETFTVRKISSGVGVERTFPLHPKIEIKVRRGKVRRAKLYL
RSLRGKAIRQIR"
complement(10910..13513)
/locus_tag="SE0917"
complement(10910..13513)
/locus_tag="SE0917"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AA004514.1"
/db_xref="GI:27315380"
/translation="MKNIMNPKIKRFTLIVLSCILSFILYGPFLYQFIKGVFSG
SGDFROMPPOMELYEHFTSLQKGYDASFGLDGYKSLAYVYSLPMLWINSMTGW
ILEOTINNVNDISQWTLQIMAVRTVITFESFVSLRLKPAEMFATILYM
STVYTNFTSWFVGNLLMLPMSLMAIERFFKERRKIGWFIATYILFNSPFYSYE
AIVGFIYRFAIPHEKDIVRWOKLYILVCATLLSVLSIYGLYTGVSFPLDNR
QNPFKITFTNLFETNNIFADGFYITISFIALFCFKLYQHYIKYKLFATATWIL
LYGFIATLSERSVKMFVALILVLFIFIKYKSLTRTSMVLLVFLAQOVLMT
IYHNGTSLYSIFDGLKYDQTLQINMPVDKNSTYRLNNRANLMSLMDVQDLR
HPDNLMPYGFKKELITDKDQWTHSVNTINYPSTHITNKIKYDARKLKSPLDREQM
LKVYLNHKSQANTDFKPNLLSNQNLNHNAMIDSKHLKVKHNGGVTLNIPRNI
VKNYQDLEIENDLSPDKHKGVNEYSQERNLSKYKRFVSPVTRAKASNLQNL
IKMGVYRVFKVGIYGNQYTLKASQOLQPVKVKERNGFTIHKKHGHLVLP
VYAGSHAMANGKPLKQVQNGIMTIPVKGSGTKIKLSYTPPYFLLITVSCGILL
SILTHYVKKR"
complement(13506..15779)

gene	Query Match	56.1%	Score 412.8;	DB 1;	Length 300892;
CDS	Best Local Similarity	74.1%	Pred. No. 3e-66;		
	Matches	522;	Conservative	0;	Mismatches 182; Indels 0; Gaps 0;
Qy	2	TTTACGTTTGTCTGATACATACATACCAACGCGACCAATCTTCAATCATTTTGTCA			
Db	208660	TTGGCGCTTAATCATCGAATACATACCTACGCGACCTACTGTTCTATGATTTTCT			
Qy	62	CGATCGTATGTTTAACTCTTTTCGTAATATCTTATATAACATCTACACATTTGTT			
Db	208720	TTATTAAGAGTTTGTAGTTTATAGTAATAACGAATATAACATCAACGACATTCGTT			

```
/organism="Staphylococcus epidermidis"
/mol_type="genomic DNA"
/strain="SRI"
/db_xref="taxon:1282"
/clone="step.1021e09"

ORIGIN
Query Match      55.2%; Score 406.4; DB 1; Length 3099;
Best Local Similarity 73.6%; Pred. No. 1.1e-64;
Matches 518; Conservative 0; Mismatches 186; Indels 0; Gaps 0;

QY 2 TTTACGTTTGTGTCATGATACATACCCACCGCAGAACTGTTCAATCATTTGTCA 61
DB 1099 TTGGCGCTTAATCATCGAATCAGATCACCCTACGCCACGTAAGTTTCTATGGATTTTCT 1158

QY 62 CGATCGTATGTTTAACTGTTTCTGTAATATCTTATATAAATCTTATATAACATCTACGACATTTGTT 121
DB 1159 TTATTAAAGGTTTGAGTTTATAGTAATAAACGAATGTAAACATCAACGACATTCGTT 1218

QY 122 TCTACTTCATATATAACCCCATACATGATTTAAATTTGTTCCGTTGCATAACATGG 181
DB 1219 TCTACTTCATATATACCCCATACGATCGAGAAATTTGTTCCGTTGCATGACGTGG 1278

QY 182 TTTTATTTTTCAGTAAAGATATAGTAATCATACTCTGTTTGTGTAATCAATTTCT 241
DB 1279 TTACGATTTTTCAGCTAAACATATAAATCGTATCTGTTTGTAGTTAAATCTAAATTGA 1338

QY 242 GCGCCATTTACCGTCACTTTAAAGGTTTCTTATCAATTTGTAATACCGTTGACATCGATA 301
DB 1339 TGCCCATTAACAGTAATCTTTAAAGCACTTTATCAATGATTAATCAATGATATCTAA 1398

QY 302 ATATCCTTTTGTGGTGACAGCTAAATTTGCGAGAAATTTGCTGTAAGTTCTTCAATA 361
DB 1399 ACATCTTTATCTGGCTGTCTGCGCAATACCGCTCTTATCTTGGAGCAATTTCTCTATA 1458

QY 362 TCAACGGCTTAATCATATATAATCGTCTGACCGTAATCAAGCCCGCAGCACTTTGTCTATAC 421
DB 1459 TCAAGGGTTTACAAATGATGATCATCTGCCCATAGTCCCAACCCAGCTACTTTATCATAT 1518

QY 422 GTATCACTTTTTCGCTGTAATTAATGATAGGTAGATTGTTGTTGTTCTAAATTTTGGCA 481
DB 1519 GTCTCGCTTTTTCAGTAATGATGATAATGAGTAGTTGTTTGTGAGCAATTTGCTTA 1578

QY 482 CAAATTTCTAAGCCATTAATGACGCAACATTAATCAATCAATGATTAATCAATGATAA 541
DB 1579 CAAATTTCTAGACCAATTTATGTTTGAAGCAATTAAGTCTAAATATATAAATCATAGGC 1638

QY 542 TGGCTAAGCGCTTTATCTAAACCGTCTTGTCCATCTACTCTGTGCCAATTTGTAATTT 601
DB 1639 TTAGATTAATGCTTTATCCAAACCCACCTTTCCATCATCTCTATATCAACAGTATAATTT 1698

QY 602 TCATGTGTAGTTTCCAAATCAAGAAATCTTGTCAAGTTTGTGTTCTCTCTCTCTCTATAA 661
DB 1699 TCATGAGTTAACTCAAGCTCTATAATCTAGCAAGATTTTGTCTCACTCTCTCTCTATAA 1758

QY 662 ATTTGGCTGATTTGTACACCTCATATACGACITTTTCTTAATAA 705
DB 1759 ATATTGTCTATATTGACCTCAGCTACATCTTAAACAATAA 1802

RESULT 14
AX145064
LOCUS
DEFINITION
Sequence 3786 from Patent WO0134809.
ACCESSION
AX145064
VERSION
AX145064.1 GI:14283629
KEYWORDS
synthetic construct
SOURCE
synthetic construct
artificial sequences.
ORGANISM
Staphylococcus epidermidis
REFERENCE
1
AUTHORS
Klimmerly, W.J.
TITLE
Staphylococcus epidermidis nucleic acids and proteins

JOURNAL Patent: WO 0134809-A 3786 17-MAY-2001;
GLAXO GROUP LIMITED (GB)
FEATURES
source
Location/Qualifiers
1..3099
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="synthetic nucleic acid sequence"

ORIGIN
Query Match      55.2%; Score 406.4; DB 6; Length 3099;
Best Local Similarity 73.6%; Pred. No. 1.1e-64;
Matches 518; Conservative 0; Mismatches 186; Indels 0; Gaps 0;

QY 2 TTTACGTTTGTGTCATGATACATACCCACCGCAGAACTGTTCAATCATTTGTCA 61
DB 1099 TTGGCGCTTTAATCATCGAATCAGATCACCCTACGCCACGTAAGTTTCTATGGATTTTCT 1158

QY 62 CGATCGTATGTTTAACTGTTTCTGTAATATCTTATATAAATCTTATATAACATCTACGACATTTGTT 121
DB 1159 TTATTAAAGGTTTGAGTTTATAGTAATAAACGAATGTAAACATCAACGACATTCGTT 1218

QY 122 TCTACTTCATATATAACCCCATACATGATTTAAATTTGTTCCGTTGCATAACATGG 181
DB 1219 TCTACTTCATATATACCCCATACGATCGAGAAATTTGTTCCGTTGCATGACGTGG 1278

QY 182 TTTTATTTTTCAGTAAAGATATAGTAATCATACTCTGTTTGTGTAATCAATTTCT 241
DB 1279 TTACGATTTTTCAGCTAAACATATAAATCGTATCTGTTTGTAGTTAAATCTAAATTGA 1338

QY 242 GCGCCATTTACCGTCACTTTAAAGGTTTCTTATCAATTTGTAATACCGTTGACATCGATA 301
DB 1339 TGCCCATTAACAGTAATCTTTAAAGCACTTTATCAATGATTAATCAATGATATCTAA 1398

QY 302 ATATCCTTTTGTGGTGACAGCTAAATTTGCGAGAAATTTGCTGTAAGTTCTTCAATA 361
DB 1399 ACATCTTTATCTGGCTGTCTGCGCAATACCGCTCTTATCTTGGAGCAATTTCTCTATA 1458

QY 362 TCAACGGCTTAATCATATATAATCGTCTGACCGTAATCAAGCCCGCAGCACTTTGTCTATAC 421
DB 1459 TCAAGGGTTTACAAATGATGATCATCTGCCCATAGTCCCAACCCAGCTACTTTATCATAT 1518

QY 422 GTATCACTTTTTCGCTGTAATTAATGATAGGTAGATTGTTGTTGTTCTAAATTTTGGCA 481
DB 1519 GTCTCGCTTTTTCAGTAATGATGATAATGAGTAGTTGTTTGTGAGCAATTTGCTTA 1578

QY 482 CAAATTTCTAAGCCATTAATGACGCAACATTAATCAATCAATGATTAATCAATGATAA 541
DB 1579 CAAATTTCTAGACCAATTTATGTTTGAAGCAATTAAGTCTAAATATATAAATCATAGGC 1638

QY 542 TGGCTAAGCGCTTTATCTAAACCGTCTTGTCCATCTACTCTGTGCCAATTTGTAATTT 601
DB 1639 TTAGATTAATGCTTTATCCAAACCCACCTTTCCATCATCTCTATATCAACAGTATAATTT 1698

QY 602 TCATGTGTAGTTTCCAAATCAAGAAATCTTGTCAAGTTTGTGTTCTCTCTCTCTCTATAA 661
DB 1699 TCATGAGTTAACTCAAGCTCTATAATCTAGCAAGATTTTGTCTCACTCTCTCTCTATAA 1758

QY 662 ATTTGGCTGATTTGTACACCTCATATACGACITTTTCTTAATAA 705
DB 1759 ATATTGTCTATATTGACCTCAGCTACATCTTAAACAATAA 1802

RESULT 15
AF269630/c
LOCUS
DEFINITION
Staphylococcus epidermidis strain SRI clone step.1014f03 genomic
sequence.
ACCESSION
AF269630
VERSION
AF269630.1 GI:9623526
KEYWORDS
Staphylococcus epidermidis
SOURCE
Staphylococcus epidermidis
ORGANISM
Staphylococcus epidermidis
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 5, 2004, 02:32:01 ; Search time 2044.61 Seconds
(without alignments)
10749.507 Million cell updates/sec

Title: US-09-006-627-3

Perfect score: 736

Sequence: 1 ATTACGTTTGTATCTA.....TTAAATTTGTATTTTCTA 736

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:

1: em_estba:*
2: em_estum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pin:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	91	12.4	1252	29	CG755566
c 2	82.2	11.2	886	29	AL146947 Anopheles
3	74.4	10.1	859	13	CA092885 SCCCL100
c 4	72.8	9.9	660	14	CD442133 EL01N0405

ALIGNMENTS

RESULT 1	CG755566	P051-1-F01.ya Ppa EcoRI BAC Library	1252 bp	DNA	linear	GSS 24-OCT-2003
LOCUS	CG755566	genomic survey sequence.				
DEFINITION	CG755566	GSS.				
ACCESSION	CG755566.1	GI:37982189				
VERSION	CG755566.1	GI:37982189				
KEYWORDS	CG755566.1	GI:37982189				
SOURCE	CG755566.1	GI:37982189				
ORGANISM	CG755566.1	GI:37982189				
REFERENCE	CG755566	P051-1-F01.ya Ppa EcoRI BAC Library	1252 bp	DNA	linear	GSS 24-OCT-2003
AUTHORS	CG755566	genomic survey sequence.				
TITLE	CG755566	GSS.				
JOURNAL	CG755566	GSS.				
MEDLINE	CG755566	GSS.				
PUBMED	CG755566	GSS.				
COMMENT	CG755566	GSS.				

5	71.2	9.7	776	28	BH373793	BH373793
6	67.2	9.1	761	28	BH386001	BH386001
7	67.2	9.1	762	28	BH379110	BH379110
8	65.8	8.9	1101	29	CNS0039G	CNS0039G
c 9	64	8.7	1302	29	BH770345	BH770345
c 10	63.2	8.6	1200	13	EX437758	EX437758
c 11	62.2	8.5	560	28	BH375414	BH375414
c 12	62.2	8.5	667	12	B1263997	B1263997
c 13	61.2	8.3	1722	11	AK002612	AK002612
c 14	60	8.2	1098	13	BX377526	BX377526
c 15	59.8	8.1	1201	13	BX360615	BX360615
c 16	59.2	8.0	583	28	BH370279	BH370279
c 17	59.2	8.0	1044	13	EX415231	EX415231
c 18	59	8.0	1056	13	EX415058	EX415058
c 19	58.2	7.9	500	28	BH393267	BH393267
c 20	57.2	7.8	705	28	AF267506	AF267506
c 21	57	7.7	715	28	AZ301083	AZ301083
c 22	56.8	7.7	1187	28	B11102	B11102
c 23	55.8	7.6	1201	13	BX355654	BX355654
c 24	55.6	7.6	1166	28	BZ549106	BZ549106
c 25	55.4	7.5	570	28	AQ989880	AQ989880
c 26	55.2	7.5	963	29	CNS00A4L	CNS00A4L
c 27	55	7.5	601	12	BM212834	BM212834
c 28	55	7.5	994	13	EX414650	EX414650
c 29	54.8	7.4	1101	29	CNS00EVL	CNS00EVL
c 30	54.6	7.4	901	28	AF075846	AF075846
c 31	54.4	7.4	922	28	AZ530260	AZ530260
c 32	54	7.3	924	28	BZ549107	BZ549107
c 33	54	7.3	1069	28	BZ559302	BZ559302
c 34	54	7.3	1392	29	CG757503	CG757503
c 35	53.8	7.3	1031	29	CNS00CF2	CNS00CF2
c 36	53.8	7.3	1101	29	CNS00HSG	CNS00HSG
c 37	53.2	7.3	1101	29	CNS0039L	CNS0039L
c 38	53.2	7.2	1092	29	CNS020K7	CNS020K7
c 39	53.2	7.2	1200	13	EX414560	EX414560
c 40	52.8	7.2	568	13	BU696722	BU696722
c 41	52.8	7.2	1236	28	BZ575791	BZ575791
c 42	52.6	7.1	877	28	BH148826	BH148826
c 43	52.6	7.1	880	28	BH155561	BH155561
c 44	52.6	7.1	890	28	AZ694975	AZ694975
c 45	52.4	7.1	994	29	CNS04NOJ	CNS04NOJ

CG755566 1252 bp DNA linear GSS 24-OCT-2003
P051-1-F01.ya Ppa EcoRI BAC Library Pristionchus pacificus genomic,
genomic survey sequence.
CG755566 GI:37982189
GSS.
Pristionchus pacificus
Pristionchus pacificus
Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
Neodiplogasteridae; Pristionchus.
Srinivasan,J., Sinz,W., Jesse,T., Wiggers-Perebolte,L., Jansen,K.,
Buntjer,J., van der Meulen,M. and Sommer,R.J.
An integrated physical and genetic map of the nematode Pristionchus
pacificus
Mol. Genet. Genomics 269 (5), 715-722 (2003)
22835951
12884007
Contact: Sommer RJ
Evolutionary Biology
Max-Planck-Institute for Developmental Biology
Spemannstr. 37-39, Tuebingen D-72076, Germany
Tel: 00497071601371
Fax: 00497071601498
Email: ralf.sommer@tuebingen.mpg.de

Class: BAC ends.		Location/Qualifiers		Location/Qualifiers						
FEATURES source	ORIGIN	Query Match Best Local Similarity Matches	12.4%; 51.1%; 214;	Score 91; 51.1%; 0;	DB 29; 1.6e-09; 0;					
1. .1252										
/organism="Pristionchus pacificus"										
/mol_type="genomic DNA"										
/strain="California"										
/db_xref="taxon:54126"										
/clone_lib="Ppa EcoRI BAC Library"										
/note="The library was generated by a partial digest of the genomic DNA with EcoRI and cloning into the BAC vector."										
1. .886										
/organism="Anopheles gambiae"										
/mol_type="genomic DNA"										
/strain="PEST"										
/db_xref="taxon:7165"										
/clone_lib="13F15"										
/clone_lib="NotreDame1"										
/note="end : SP6"										
1. .886										
/organism="Anopheles gambiae"										
/mol_type="genomic DNA"										
/strain="PEST"										
/db_xref="taxon:7165"										
/clone_lib="13F15"										
/clone_lib="NotreDame1"										
/note="end : SP6"										
1. .886										
/organism="Anopheles gambiae"										
/mol_type="genomic DNA"										
/strain="PEST"										
/db_xref="taxon:7165"										
/clone_lib="13F15"										
/clone_lib="NotreDame1"										
/note="end : SP6"										
1. .886										
/organism="Anopheles gambiae"										
/mol_type="genomic DNA"										
/strain="PEST"										
/db_xref="taxon:7165"										
/clone_lib="13F15"										
/clone_lib="NotreDame1"										
/note="end : SP6"										
1. .886										
/organism="Anopheles gambiae"										
/mol_type="genomic DNA"										
/strain="PEST"										
/db_xref="taxon:7165"										
/clone_lib="13F15"										
/clone_lib="NotreDame1"										
/note="end : SP6"										
1. .886										
/organism="Anopheles gambiae"										
/mol_type="genomic DNA"										
/strain="PEST"										
/db_xref="taxon:7165"										
/clone_lib="13F15"										
/clone_lib="NotreDame1"										


```

DEFINITION AG-ND-144F15,TF ND-TAM Anopheles gambiae genomic clone
ACCESSION AG-ND-144F15, genomic survey sequence.
VERSION BH373793.1 GI:17319935
KEYWORDS GSS.
SOURCE Anopheles gambiae (African malaria mosquito)
ORGANISM Anopheles gambiae
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea;
Anopheles.
REFERENCE 1 (bases 1 to 776)
AUTHORS Hong,Y.S., Hogan,J.R., Wang,X., Sarkar,A., Sim,C., Loftus,B.J.,
Ren,C., Huff,E.R., Carlile,J.L., Black,K., Zhang,H.-B.,
Gardner,M.J. and Collins,F.H.
TITLE Construction of a BAC library and generation of BAC end
sequence-tagged connectors for genome sequencing of the African
malaria mosquito Anopheles gambiae
JOURNAL Mol. Genet. Genomics 268 (6), 720-728 (2003)
MEDLINE 22542063
PUBMED 12655398
COMMENT Other_GSSs: AG-ND-144F15.TR
Contact: Brendan J Loftus
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0208
Fax: 301 838 3543
Email: bjlloftus@tigr.org
This clone is from an A. gambiae BAC library (ND-TAM) provided by
F.H. Collins and sequenced by The Institute for Genomic Research
(TIGR). The BAC library was generated from A. gambiae PEST strain
DNA. All DNA was extracted from newly hatched first instar larvae
to minimize the inclusion of DNA from microorganisms that inhabit
the gut. The DNA is derived from mixed sexes of larvae. The BAC
library was constructed at Texas A&M University BAC Center
University, College Station, Texas 77843-2123, USA using a HindIII
Seq primer: M13 For
Class: BAC ends.
FEATURES             source
    source
        Location/Qualifiers
            1..776
                /organism="Anopheles gambiae"
                /mol_type="genomic DNA"
                /strain="PEST"
                /db_xref="taxon:7165"
                /clone="AG-ND-144F15"
                /clone_lib="ND-TAM"
                /note="Vector: pECBAC1; Site_1: HindIII"
ORIGIN
Query Match          9.7%; Score 71.2; DB 28; Length 776;
Best Local Similarity 51.0%; Pred. No. 2.9e-05;
Matches 250; Conservative 0; Mismatches 228; Indels 12; Gaps 3;
QY 23 ACATACCAACGCGCACGAACTGTTTCAATCATTTTGTTCACGATCGTATGTTTAACTTG 82
Db 277 ATATAACCCATTCGACGACAGTGATCATGTTGCTTGAAAGCCCTTATCTACTTTT 336
QY 83 TTTCGTAATATCTTATATAACATCTACGACATTTGTTTCTTCTACTTCACTATTATAACC 142
Db 337 GTCTTACATAGTTAATATAGACATCTATAGATTAGTCCTCTGTTAAAGCCCAATCCCC 396
QY 143 CATACATGATTTAAATTTGTTCCGTTGCATACATGTTTATTTATTTTCAGCTAGAAGA 202
Db 397 CAAACTGCTTCTGCAATTTTTCACGCGTTTATAACCCCTGTTTATTTTACCAGAAGATAT 456
QY 203 TATAGTAATCATACTCTGTTTTTTGTTAAATTCATTTT-----CTGCGCCATTTACCGTC 256
Db 457 TCTAACAGGCGATATTTCTTTACCGTCAGAGTAATCTCTGCCCTCGCCCTTTTACAAAC 516
QY 257 ACTTTAAAGCGTTCTTATCAATTTGTAAT----ACCGTTGACATCGATAATATCTTTTG 312
Db 517 TTCTATAGCAATCCAATTCCAAGATCATCTGCAGTATAGTTAGTACCGGAAGGCTATT 576

```

```

RESULT 6
LOCUS BH386001
DEFINITION AG-ND-12113.TF ND-TAM Anopheles gambiae genomic clone AG-ND-12113,
Genomic survey sequence.
ACCESSION BH386001
VERSION BH386001.1 GI:17332143
KEYWORDS GSS.
SOURCE Anopheles gambiae (African malaria mosquito)
ORGANISM Anopheles gambiae
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea;
Anopheles.
REFERENCE 1 (bases 1 to 761)
AUTHORS Hong,Y.S., Hogan,J.R., Wang,X., Sarkar,A., Sim,C., Loftus,B.J.,
Ren,C., Huff,E.R., Carlile,J.L., Black,K., Zhang,H.-B.,
Gardner,M.J. and Collins,F.H.
TITLE Construction of a BAC library and generation of BAC end
sequence-tagged connectors for genome sequencing of the African
malaria mosquito Anopheles gambiae
JOURNAL Mol. Genet. Genomics 268 (6), 720-728 (2003)
MEDLINE 22542063
PUBMED 12655398
COMMENT Other_GSSs: AG-ND-12113.TR
Contact: Brendan J Loftus
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0208
Fax: 301 838 3543
Email: bjlloftus@tigr.org
This clone is from an A. gambiae BAC library (ND-TAM) provided by
F.H. Collins and sequenced by The Institute for Genomic Research
(TIGR). The BAC library was generated from A. gambiae PEST strain
DNA. All DNA was extracted from newly hatched first instar larvae
to minimize the inclusion of DNA from microorganisms that inhabit
the gut. The DNA is derived from mixed sexes of larvae. The BAC
library was constructed at Texas A&M University BAC Center
University, College Station, Texas 77843-2123, USA using a HindIII
Seq primer: M13 For
Class: BAC ends.
FEATURES             source
    source
        Location/Qualifiers
            1..761
                /organism="Anopheles gambiae"
                /mol_type="genomic DNA"
                /strain="PEST"
                /db_xref="taxon:7165"
                /clone="AG-ND-12113"
                /clone_lib="ND-TAM"
                /note="Vector: pECBAC1; Site_1: HindIII"
ORIGIN
Query Match          9.1%; Score 67.2; DB 28; Length 761;
Best Local Similarity 51.7%; Pred. No. 0.00021;
QY 313 TGCTGACGACGTAATAATTGACGATTTCTGTAAAGTTCTTCAATATCAACGGCTT 372
Db 577 CTGATTTCTTCGGCTTAGCGCATTAATCTCCGAGTAATCTTTCGAATGAATGGCTT 636
QY 373 AACTATATATCGTCTGCACCGCTTAATCAAGCCAGCCCACTTTGTCTACATGATCACTTT 432
Db 637 GGTAAGATAGTATCTGCGCCCTTATCAAGCCCTTTAGCTTGTCTTGTATGCTAGAGAG 696
QY 433 CGCTGTAATTAATGATAGAG--TGTAGATTGTTGTTGTTCTAATTTTGGACAAATTTCT 490
Db 697 TGGCGTCAGCATTAATAAACTGGAATTTCCGGCTTAACGCGCTTTATTTTCAACATATTCA 756
QY 491 AAGCCATTAA 500
Db 757 AAGCCTTTCA 766

```

Matches 178; Conservative 0; Mismatches 163; Indels 3; Gaps 1;

Qy 324 GTAAATGCGACGAAATCTTCTGCTAAAGGTTCTTCAATATCAACCGCTTAACATATAAT 383
 Db 110 GAAGGATTCCTTAATTTTGTAAAGCAATAATTCAGTATCGAAGGTTTTGTGATAAT 169

Qy 384 CGTCTGCAACGTAATCAAGCCAGCAACTTTGTCTATACGTATCACTTTTCGCTGTAATTA 443
 Db 170 CATCGCTCCGATTTGGTACCCCTTTAAGGATGTCTTCACGCGAGGTTTTTAGCGGTAAGGA 229

Qy 444 TAATGATAGGTAGATTGTTG---TTGCTAAATTTTCCGACAAATTTCTAAGCCATTAA 500
 Db 230 AGATAATCGGAATATTTTATCAATTTTTTAAATTTCTTCAGCTAAAGAGATCCATCCT 289

Qy 501 TTGACGGCAACATTAATCTAATGATTAATGATTAATCATAGTAAATGCTAAGCCCTTTATCTA 560
 Db 290 TTTTGGCATCATGACATCGGTAAATGCAAAATGTCAAATTCATTTTCGGTAAATTTCTTTA 349

Qy 561 AACGCTCTTGTCATCATCTCTGTGTCACCATTTGTTGTCATCTTCTACTATTAAATTT 620
 Db 350 GTCCCTTGTTCCATCTCGGCGACGAGTTACATCGAAGTTAATTAATGTTAAATATCCT 409

Qy 621 CAAGAAATCTTGCTAAGTTTGTTCATCTTCTTCTACTATTAAATTT 664
 Db 410 TCAGTACAGCTCCGAAGCTTTGGTCTCTTCTACGAGTAAGATT 453

RESULT 7
 BH379110
 LOCUS
 DEFINITION AG-ND-167112.TF ND-TAM Anopheles gambiae genomic clone
 AG-ND-167112, genomic survey sequence.

ACCESSION BH379110
 VERSION BH379110.1 GI:17325252

KEYWORDS GSS.
 SOURCE Anopheles gambiae (African malaria mosquito)
 ORGANISM Anopheles gambiae
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 Neoptera; Endopterygota; Diptera; Nematocera; Culicoides;
 Anopheles.

REFERENCE 1 (bases 1 to 762)
 AUTHORS Hong, Y.S., Hogan, J.R., Wang, X., Sarkar, A., Sim, C., Loftus, B.J.,
 Ren, C., Huff, E.R., Carlile, J.L., Black, K., Zhang, H.-B.,
 Gardner, M.J. and Collins, F.H.

TITLE Construction of a BAC library and generation of BAC end
 sequence-tagged connectors for genome sequencing of the African
 malaria mosquito Anopheles gambiae

JOURNAL Mol. Genet. Genomics 268 (6), 720-728 (2003)
 MEDLINE 22542063
 PUBMED 12653398

COMMENT Other GSSs: AG-ND-167112.TR
 Contact: Brendan J Loftus
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0208
 Fax: 301 838 3543
 Email: bjlloftus@igr.org

This clone is from an A. gambiae BAC library (ND-TAM) provided by
 F.H. Collins and sequenced by The Institute for Genomic Research
 (TIGR). The BAC library was generated from A. gambiae PEST strain
 DNA. All DNA was extracted from newly hatched first instar larvae
 to minimize the inclusion of DNA from microorganisms that inhabit
 the gut. The DNA is derived from mixed sexes of larvae. The BAC
 library was constructed at Texas A&M University BAC Center
 University, College Station, Texas 77843-2123, USA using a HindIII
 partial digest.

Seq primer: M13 For

Class: BAC ends.

Location/Qualifiers

1..762

/organism="Anopheles gambiae"

/mol_type="genomic DNA"

FEATURES

source

1..1101

Location/Qualifiers

/strain="PEST"
 /db_xref="taxon:7165"
 /clone="AG-ND-167112"
 /clone_lib="ND-TAM"
 /note="Vector: pECBAC1; Site_1: HindIII"

Query Match 9.1%; Score 67.2; DB 28; Length 762;
 Best Local Similarity 51.7%; Pred. No. 0.00021;

Matches 178; Conservative 0; Mismatches 163; Indels 3; Gaps 1;

Qy 324 GTAAATGCGACGAAATCTTCTGCTAAAGGTTCTTCAATATCAACCGCTTAACATATAAT 383
 Db 110 GAAGGATTCCTTAATTTTGTAAAGCAATAATTCAGTATCGAAGGTTTTGTGATAAT 169

Qy 384 CGTCTGCAACGTAATCAAGCCAGCAACTTTGTCTATACGTATCACTTTTCGCTGTAATTA 443
 Db 170 CATCGCTCCGATTTGGTACCCCTTTAAGGATGTCTTCACGCGAGGTTTTTAGCGGTAAGGA 229

Qy 444 TAATGATAGGTAGATTGTTG---TTGCTAAATTTTCCGACAAATTTCTAAGCCATTAA 500
 Db 230 AGATAATCGGAATATTTTATCAATTTTTTAAATTTCTTCAGCTAAAGAGATCCATCCT 289

Qy 501 TTGACGGCAACATTAATCTAATGATTAATGATTAATCATAGTAAATGCTAAGCCCTTTATCTA 560
 Db 290 TTTTGGCATCATGACATCGGTAAATGCAAAATGTCAAATTCATTTTCGGTAAATTTCTTTA 349

Qy 561 AACGCTCTTGTCATCATCTCTGTGTCACCATTTGTTGTCATCTTCTACTATTAAATTT 620
 Db 350 GTCCCTTGTTCCATCTCGGCGACGAGTTACATCGAAGTTAATTAATGTTAAATATCCT 409

Qy 621 CAAGAAATCTTGCTAAGTTTGTTCATCTTCTTCTACTATTAAATTT 664
 Db 410 TCAGTACAGCTCCGAAGCTTTGGTCTCTTCTACGAGTAAGATT 453

RESULT 8

CNS0039G

LOCUS

DEFINITION

Drosophila melanogaster genome survey sequence TET3 end of BAC #

BACR08K10 of RPCI-98 library from Drosophila melanogaster (fruit

fly), genomic survey sequence.

AL063921

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Drosophila melanogaster (fruit fly)

Drosophila melanogaster

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

Ephydroidea; Drosophilidae; Drosophila.

1 (bases 1 to 1101)

Genoscope.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Direct Submission

Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :

BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr

- Web : www.genoscope.cns.fr)

Determination of this BAC-end sequence was carried out as part of a

collaboration with the Berkeley Drosophila Genome Project (BDGP).

The BDGP is constructing a physical map of the Drosophila

melanogaster genome using these BACs. For further information

please see <http://www.fruitfly.org> The BDGP Drosophila

melanogaster BAC library was prepared by Kazutoyo Osoegawa and

Aaron Mamoser in Peter de Jong's laboratory in the Department of

Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,

NY. The library is named RPCI-98 and was constructed by partial

EcoRI digestion of Drosophila DNA provided by the BDGP from the

isogenic strain Y2; cn bw sp, the same strain used for the BDGP's

P1 and EST libraries. A more detailed description of the library

and how to order individual BAC clones, the entire library, or

filters for hybridization from the BACPAC Resource Center can be

found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

clone NF092E12PL 5', mRNA sequence.	AK002612/c
BI263997	LOCUS
BI263997.1 GI:14865815	DEFINITION
EST.	
Medicago truncatula (barrel medic)	ACCESSION
Medicago truncatula	VERSION
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.	KEYWORDS
	SOURCE
	ORGANISM
REFERENCE	
1 (bases 1 to 667)	
Liu, J., Scott, A.D., Harris, A.R., Gonzales, R.A., Bell, C.J., Flores, H.R., Inman, J.T., Weller, J.W., May, G.D. and Harrison, M.J.	REFERENCE
Expressed Sequence Tags from the Samuel Roberts Noble Foundation	AUTHORS
Medicago truncatula phosphate-starved leaf library	TITLE
Unpublished (2000)	JOURNAL
Contact: Harrison MJ	MEDLINE
Plant Biology Division	PUBMED
The Samuel Roberts Noble Foundation	REFERENCE
2510 Sam Noble Parkway, Ardmore, OK 73402, USA	AUTHORS
Tel: 580 221 7325	
Fax: 580 221 7380	TITLE
Email: mjharrison@noble.org	JOURNAL
Insert Length: 667 Std Error: 0.00	MEDLINE
Plate: 092 row: E column: 12	PUBMED
Seq primer: TCACACAGGAACAGCTATGAC.	REFERENCE
	AUTHORS
FEATURES	
source	
1. 667	
/organism="Medicago truncatula"	
/mol_type="mRNA"	
/db_xref="taxon:3880"	
/clone="NF092E12PL"	
/tissue_type="leaf"	
/dev_stage="trifoliolate"	
/clone_lib="Phosphate starved leaf"	
/notes="Vector: Lambda Zap; At the trifoliolate stage, M. truncatula plants were transplanted to phosphate-free sand and grown for a further 30 days. During this 30 day period, the plants were fertilized twice weekly with 1/2 Hoaglands solution containing only 20uM potassium phosphate. RNA was prepared from above ground tissues."	
	ORIGIN
Query Match 8.5%; Score 62.2; DB 12; Length 667;	
Best Local Similarity 50.4%; Pred. No. 0.0025;	
Matches 179; Conservative 0; Mismatches 173; Indels 3; Gaps 1;	
314 GGCTGACGCGTAAATTCACGAATTCCTTGCTAAAGATTCCTCAATATCAACACGGCTTA 373	Qy
468 GCCTGACGCGCAGCAGCCAGCATACGGCGCAGGTTTAAACGGTTTT 409	Db
374 ACTATATAATCGTCTCCAGCTTAATCAAGCCAGCACTTTGTCATACGTATCACTTTTC 433	Qy
408 GGAATGTAGTCGTGACGCGCAATCTCCAGGCTCAGATACGCTCCACTTCTTCCCTTTC 349	Db
434 GCTGTAAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 490	Qy
348 GCGGTACCAATATGATGCGGATCGGTTGCTCTGACTACGAAGACGTCGCGCAATCGAC 289	Db
491 AAGCCATTAATTCAGCGCAACATTAAATCTAATATGATTAATCATAGTAATGCTTAAGC 550	Qy
288 AAGCCATCTTCACGAGTAAATTAATCCAGTACCATAAGATGGAAGATTCACGAGTC 229	Db
551 GCTTATCTAAACCGTCTGTGTCATCATCACTCTGTGTCACATGTAATTTTCATGTGTG 610	Qy
228 AGCAGGCGATCCATCTGTGTTCTGCATAGCAGCGCTTCGAACCTCGAAGCGCTTGTGCGTG 169	Db
611 AGTTCCAAATTCAGAAATCTTGCTAGTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTT 665	Qy
168 AGATAACGTTCCAGCAGCGCAGCGCGCGATGCTGCTATCGACCCACGAGTAATCT 114	Db
	RESULT 13

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 5, 2004, 03:08:42 ; Search time 55.6323 Seconds
(without alignments)
7341.850 Million cell updates/sec

Title: US-09-006-627-3
Perfect score: 736
Sequence: 1 ATTTACGTTTGTGTCATCGTA.....TTAAATTTGTATTTTCTA 736

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents_NA.*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PTCUS_COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	736	100.0	736	3	US-08-771-110-1
2	736	100.0	736	4	US-09-642-000-3
3	736	100.0	2201	4	US-09-642-000-1
4	726	98.6	1440	4	US-08-956-171E-725
5	401.8	54.6	669	4	US-09-134-001C-375
6	187.2	25.4	687	4	US-09-107-532A-2882
7	149.8	20.4	741	4	US-09-134-001C-338
8	145.8	19.8	555	4	US-09-134-000C-415
9	144	19.6	684	4	US-09-328-352-4049
10	130.2	17.7	760	4	US-08-956-171E-822
11	129	17.5	696	4	US-09-134-000C-2259
12	127.4	17.3	900	2	US-08-771-455-1
13	125.8	17.1	696	4	US-09-134-000C-1740
14	125.6	17.1	702	4	US-09-082-077-2
15	125.6	17.1	3731	4	US-09-082-077-1
16	125.6	17.1	9425	4	US-08-956-171E-87
17	125	17.0	705	4	US-09-134-001C-226
18	121.8	16.5	687	4	US-09-134-001C-1246
19	120.8	16.4	711	4	US-09-107-532A-3219
20	119.6	16.2	724	3	US-08-771-321-1
21	119.6	16.2	3636	4	US-08-961-527-78
22	116.4	15.8	2400	3	US-08-963-901-1
23	116.4	15.8	2400	3	US-08-963-901-5
24	113.8	15.5	738	4	US-09-134-001C-942
25	113.8	15.5	850	2	US-08-771-783-1
26	113.4	15.4	741	4	US-09-107-532A-481
27	111	15.1	4279	4	US-08-993-825-1

28	110.2	15.0	1830121	4	US-09-557-884-1	Sequence 1, Appli
29	110.2	15.0	1830121	4	US-09-643-990A-1	Sequence 1, Appli
30	109.8	14.9	720	4	US-09-711-164-196	Sequence 196, App
31	106.4	14.5	684	4	US-09-107-532A-3115	Sequence 3115, Ap
32	106.4	14.5	1060	1	US-08-850-118-1	Sequence 1, Appli
33	106.4	14.5	1060	2	US-09-008-253-1	Sequence 1, Appli
34	106.4	14.5	1060	3	US-09-093-335-1	Sequence 1, Appli
35	106.2	14.4	9909	4	US-08-961-527-12	Sequence 12, Appli
36	105.4	14.3	726	4	US-09-107-532A-1192	Sequence 1192, Ap
37	105	14.3	705	4	US-09-107-532A-3092	Sequence 3092, Ap
38	101	13.7	729	4	US-09-634-238-133	Sequence 133, App
39	100.6	13.7	7363	4	US-08-956-171E-19	Sequence 19, Appli
40	100.6	13.6	666	4	US-09-107-532A-823	Sequence 823, App
41	100.2	13.6	666	4	US-09-134-000C-3118	Sequence 3118, Ap
42	100	13.6	1717	4	US-09-634-238-6	Sequence 6, Appli
43	98.8	13.4	702	4	US-09-328-352-1140	Sequence 1140, Ap
44	98.6	13.4	750	4	US-09-328-352-564	Sequence 564, App
45	98.6	13.4	2279	3	US-09-081-686-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-08-771-110-1
; Sequence 1, Application US/08771110
; Patent No. 6084086
; GENERAL INFORMATION:
; APPLICANT: Wallis, Nicola G.
; APPLICANT: Hodgson, John E.
; TITLE OF INVENTION: No. 6084086e1 Compounds
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Smithkline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-2799
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/771,110
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Gimmi, Edward R.
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: P31316
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 736 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-771-110-1

Query Match 100.0%; Score 736; DB 3; Length 736;
Best Local Similarity 100.0%; Pred. No. 2.8e-175;
Matches 736; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1 ATTTACGTTTGTGTCATCGTATCACATCCGACCCAGCACTGTTCAATCATTTTGTGTC 60
Db 1 ATTTACGTTTGTGTCATCGTATCACATCCGACCCAGCACTGTTCAATCATTTTGTGTC 60
Oy 61 ACGATCGTATGTTTAACTGTTTTCGTAATAATCTTATATAAACATCTACGACATTTGT 120

```
Db 61 ACAGTCTAGTGGTTTAACTGTTTGGTAATAATCTTATATAAATCTAGGACATTTGT 120
Qy 121 TTCTACTCTACTATTATAAACCACATACATGATTAAATTTGTTCCGTTGCATAACATG 180
Db 121 TTCTACTCTACTATTATAAACCACATACATGATTAAATTTGTTCCGTTGCATAACATG 180
Qy 181 GTTTTATTTTCAGCTAGAGATATAGTAATATCATCTGTTTGTGTTTAAATCAATTTTC 240
Db 181 GTTTTATTTTCAGCTAGAGATATAGTAATATCATCTGTTTGTGTTTAAATCAATTTTC 240
Qy 241 TGGGCCATTTACCGTCACTTTTAAAGGTTTCTTATCAATTTGTAATACCGTTGCATCGAT 300
Db 241 TGGGCCATTTACCGTCACTTTTAAAGGTTTCTTATCAATTTGTAATACCGTTGCATCGAT 300
Qy 301 AATATCCTTTTGGCTGAGAGATAGTAATATCAATTTGTAATACCGTTGCATCGAT 360
Db 301 AATATCCTTTTGGCTGAGAGATAGTAATATCAATTTGTAATACCGTTGCATCGAT 360
Qy 361 ATCAAAAGGCTTAACTATATTAATCGTCTGACCGTAAATCAAGCCAGCACTTTGTCATA 420
Db 361 ATCAAAAGGCTTAACTATATTAATCGTCTGACCGTAAATCAAGCCAGCACTTTGTCATA 420
Qy 421 CGTATCACTTTTTCGCTGTAATATAATGATAGGTGTAATTTGTTGTTCTAAATTTGCG 480
Db 421 CGTATCACTTTTTCGCTGTAATATAATGATAGGTGTAATTTGTTGTTCTAAATTTGCG 480
Qy 481 ACAAAATTTCAAGCCATTAATGACGCAACATTAATCAATTAATGATTAATCAATGTA 540
Db 481 ACAAAATTTCAAGCCATTAATGACGCAACATTAATCAATTAATGATTAATCAATGTA 540
Qy 541 ATGCTAAGCGCTTTATCTAAACCGTCTGTGCAATCATCTCTGTGTCACATTTGTAAT 600
Db 541 ATGCTAAGCGCTTTATCTAAACCGTCTGTGCAATCATCTCTGTGTCACATTTGTAAT 600
Qy 601 TTCAATGTTGAGTTTCCAAATCAAGAAATCTTGTGTAAGTTTGTTCATCTTCTACTATA 660
Db 601 TTCAATGTTGAGTTTCCAAATCAAGAAATCTTGTGTAAGTTTGTTCATCTTCTACTATA 660
Qy 661 AATTTGGCGTATTTGTAACCTCATATAGCACTTTTCTTAAAGGTAATATATTTTAA 720
Db 661 AATTTGGCGTATTTGTAACCTCATATAGCACTTTTCTTAAAGGTAATATATTTTAA 720
Qy 721 AATTTGTAATTTTCTTA 736
Db 721 AATTTGTAATTTTCTTA 736
```

RESULT 2

```
US-09-642-000-3
; Sequence 3, Application US/09642000
; Patent No. 6548281
; GENERAL INFORMATION:
; APPLICANT: Wallis, Nicola G.
; Shilling, Lisa K.
; Mooney, Jeffrey L.
; Debouck, Christine
; Zhong, Yi
; Jaworski, Deborah D.
; Wang, Min
; Throup, John P.
; TITLE OF INVENTION: Histidine Kinase
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dechert, Price & Rhoads
; STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103-2793
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
```

```
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/642,000
; FILING DATE: 18-Aug-2000
; CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/006,627
; FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
; NAME: Falk, Stephen T
; REGISTRATION NUMBER: 36,795
; REFERENCE/DOCKET NUMBER: GM10127
TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-994-2488
; TELEFAX: 215-994-2222
; TELEX: <Unknown>
INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 736 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-642-000-3

Query Match 100.0%; Score 736; DB 4; Length 736;
Best Local Similarity 100.0%; Pred. No. 2.8e-175;
Matches 736; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATTTAGCTTTTGTTCATCGTATACATACCCAGCCAGCACTGTTTCAATCAATTTTGTG 60
Db 1 ATTTAGCTTTTGTTCATCGTATACATACCCAGCCAGCACTGTTTCAATCAATTTTGTG 60
Qy 61 ACGATCGTATGTTTAACTGTTTCGTAAATATCTTATATAAATCTTACGACATTTGT 120
Db 61 ACGATCGTATGTTTAACTGTTTCGTAAATATCTTATATAAATCTTACGACATTTGT 120
Qy 121 TTCTACTTCACTATTAATACCCCATACATGATTTAAATTTGTTCCGTTGCATAACATG 180
Db 121 TTCTACTTCACTATTAATACCCCATACATGATTTAAATTTGTTCCGTTGCATAACATG 180
Qy 181 GTTTTATTTTTCAGCTAGAGATATAGTAAATCATACTCTGTTTGTGTTTAAATTTCA 240
Db 181 GTTTTATTTTTCAGCTAGAGATATAGTAAATCATACTCTGTTTGTGTTTAAATTTCA 240
Qy 241 TGGGCCATTTACCGTCACTTTTAAAGCGTTTCTTATCAATTTGTAATACCGTTGACATCGAT 300
Db 241 TGGGCCATTTACCGTCACTTTTAAAGCGTTTCTTATCAATTTGTAATACCGTTGACATCGAT 300
Qy 301 AATATCCTTTTGGCTGAGAGATAGTAATTTGACGAATTTGTCGAATTTCTTCAAT 360
Db 301 AATATCCTTTTGGCTGAGAGATAGTAATTTGACGAATTTGTCGAATTTCTTCAAT 360
Qy 361 ATCAAAAGGCTTAACTATATATCGTCTGACCGTAAATCAAGCCAGCACTTTGTCATA 420
Db 361 ATCAAAAGGCTTAACTATATATCGTCTGACCGTAAATCAAGCCAGCACTTTGTCATA 420
Qy 421 CGTATCACTTTTTCGCTGTAATATAATGATAGGTGTAATTTGTTGTTCTAAATTTGCG 480
Db 421 CGTATCACTTTTTCGCTGTAATATAATGATAGGTGTAATTTGTTGTTCTAAATTTGCG 480
Qy 481 ACAAAATTTCAAGCCATTAATGACGCAACATTAATCAATTAATGATTAATCAATGTA 540
Db 481 ACAAAATTTCAAGCCATTAATGACGCAACATTAATCAATTAATGATTAATCAATGTA 540
Qy 541 ATGCTAAGCGCTTTATCTAAACCGTCTGTGCAATCATCTCTGTGTCACATTTGTAAT 600
Db 541 ATGCTAAGCGCTTTATCTAAACCGTCTGTGCAATCATCTCTGTGTCACATTTGTAAT 600
Qy 601 TTCAATGTTGAGTTTCCAAATCAAGAAATCTTGTGTAAGTTTGTTCATCTTCTACTATA 660
Db 601 TTCAATGTTGAGTTTCCAAATCAAGAAATCTTGTGTAAGTTTGTTCATCTTCTACTATA 660
```

QY 661 AATTGCGTCATTGTACACCTCATATACGACTTTTCTAATAAGGTAATATTTTAA 720
Db 661 AATTGCGTCATTGTACACCTCATATACGACTTTTCTAATAAGGTAATATTTTAA 720
QY 721 ATTTTGTATTTTCTA 736
Db 721 ATTTTGTATTTTCTA 736

RESULT 3

US-09-642-000-1/c
; Sequence 1, Application US/09642000
; Patent No. 6548281
; GENERAL INFORMATION:
; APPLICANT: Wallis, Nicola G.
; Shilling, Lisa K.
; Mooney, Jeffrey L.
; Debouck, Christine
; Zhong, Yiyi
; Jaworski, Deborah D.
; Wang, Min
; Throup, John P.
; TITLE OF INVENTION: Histidine Kinase
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESS: Dechert, Price & Rhoads
; STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103-2793

COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/642,000
; FILING DATE: 18-Aug-2000
; CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/006,627
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Falk, Stephen T
; REGISTRATION NUMBER: 36,795
; REFERENCE/DOCKET NUMBER: GM10127
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-994-2488
; TELEFAX: 215-994-2222
; TELEX: <Unknown>

INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2201 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:

US-09-642-000-1
Query Match 100.0%; Score 736; DB 4; Length 2201;
Best Local Similarity 100.0%; Pred. No. 3.9e-175;
Matches 736; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATTTAGTTTGTCTATCATATACCCAGCCAGCACTGTTTCAATCATTTTGTGTC 60
Db 759 ATTTAGTTTGTCTATCATATACCCAGCCAGCACTGTTTCAATCATTTTGTGTC 700
QY 61 ACGATCGTATGTTTAACTGTTTTCGTAATATCTTATATAAATCTACGACATTTGT 120
Db 699 ACGATCGTATGTTTAACTGTTTTCGTAATATCTTATATAAATCTACGACATTTGT 640

QY 121 TTCTACTTCACTATTATAAACCCTTACATGATTTAAATTTTGTCCCGTGCATAACATG 180
Db 639 TTCTACTTCACTATTATAAACCCTTACATGATTTAAATTTTGTCCCGTGCATAACATG 580
QY 181 GTTTTATTTTTCAGCTAGAGATATAGTAATATCATCTGTTTGTGTTTGTGTTTCAATTC 240
Db 579 GTTTTATTTTTCAGCTAGAGATATAGTAATATCATCTGTTTGTGTTTGTGTTTCAATTC 520
QY 241 TGGCCATTTAACCGTCACCTTTAAAAGCGTCTTATCAATTTGATTAACCGTTGACATCGAT 300
Db 519 TGGCCATTTAACCGTCACCTTTAAAAGCGTCTTATCAATTTGATTAACCGTTGACATCGAT 460
QY 301 AATATCCTTTTGTGCTGACGACGATAAATGCAAGAAATCTTGTCTAAAAGTCTTCAAT 360
Db 459 AATATCCTTTTGTGCTGACGACGATAAATGCAAGAAATCTTGTCTAAAAGTCTTCAAT 400
QY 361 ATCAAGCGCTTAACTATATATATCGTCTGACCGTAAATCAAGCCGAGCACTTTGTGTCATA 420
Db 399 ATCAAGCGCTTAACTATATATCGTCTGACCGTAAATCAAGCCGAGCACTTTGTGTCATA 340
QY 421 CGTATCACTTTTTCGCTGTAATTAATGATAGGTGTAGATTGTTGTTCTAAATTTTTCG 480
Db 339 CGTATCACTTTTTCGCTGTAATTAATGATAGGTGTAGATTGTTGTTCTAAATTTTTCG 280
QY 481 ACAAATTTCTAAGCCATTAAATGACGGCAACATTAATCTAATATGATTAATATCATAGTA 540
Db 279 ACAAATTTCTAAGCCATTAAATGACGGCAACATTAATCTAATATGATTAATATCATAGTA 220
QY 541 ATGGCTAAGCGCTTTATCTAAACCGTCTTGTGCAATCATCTGTTGTCACATGTTAAT 600
Db 219 ATGGCTAAGCGCTTTATCTAAACCGTCTTGTGCAATCATCTGTTGTCACATGTTAAT 160
QY 601 TTCAATGTGTAGTTTCCAAATTCAGAAATCTTGTGTAAGTTTGTGTTCTACTATTAA 660
Db 159 TTCAATGTGTAGTTTCCAAATTCAGAAATCTTGTGTAAGTTTGTGTTCTACTATTAA 100
QY 661 AATTGCGTCATTGTACACCTCATATATACGACTTTTCTAATAAGGTAATATTTTAA 720
Db 99 AATTGCGTCATTGTACACCTCATATATACGACTTTTCTAATAAGGTAATATTTTAA 40
QY 721 ATTTTGTATTTTCTA 736
Db 39 ATTTTGTATTTTCTA 24

RESULT 4

US-08-956-171E-725/c
; Sequence 725, Application US/08956171E
; Patent No. 6593114
; GENERAL INFORMATION:

APPLICANT: Charles Kunsch
; Gil H. Choi
; Patrick S. Dillon
; Craig A. Rosen
; Steven C. Barash
; Michael R. Fannon

TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5256
; CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850

COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: Hp Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/956,171E
; FILING DATE: 20-Oct-1997

CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/009,861
FILING DATE: January 5, 1996
APPLICATION NUMBER: 08/781,986
FILING DATE: January 3, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Mark J. Hyman
REGISTRATION NUMBER: 46,789
REFERENCE/DOCKET NUMBER: PB248P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (240) 314-1224
TELEFAX: (301) 309-8439
INFORMATION FOR SEQ ID NO: 725:
SEQUENCE CHARACTERISTICS:
LENGTH: 1440 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 725:
US-08-956-171B-725

Query Match 98.6%; Score 726; DB 4; Length 1440;
Best Local Similarity 99.3%; Pred. No. 1.1e-172;
Matches 726; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

```
QY 1 ATTACGTTTGTGTCATCGTATCACATACCCAGCCAGCAACTGTTTCAATCATTTTGTGTC 60
DB 734 ATTACGTTTGTGTCATCGTATCACATACCCAGCCAGCAACTGTTTCAATCATTTTGTGTC 675
QY 61 ACGATCGTATGTTTAACTGTTTGGTAAATATCTTATATAACATCTACGACATTTGT 120
DB 674 ACGATCGTATGTTTAACTGTTTGGTAAATATCTTATATAACATCTACGACATTTGT 615
QY 121 TTCTACTTCACTATTAACCCCATACATGATTTAAATTTGTTCCGTTGCAATACATG 180
DB 614 TTCTACTTCACTATTAACCCCATACATGATTTAAATTTGTTCCGTTGCAATACATG 555
QY 181 GTTTTATTTTTCAGTAGAGATATAGTAAATCATATCTCTGTTTGTGTTAAATTCATTTTC 240
DB 554 GTTTTATTTTTCAGTAGAGATATAGTAAATCATATCTCTGTTTGTGTTAAATTCATTTTC 495
QY 241 TGGGCCATTTACCGTCACTTTTAAAGCGTTCTTATCAATTTGTAATACCGTTGACATCGAT 300
DB 494 TGGGCCATTTACCGTCACTTTTAAAGCGTTCTTATCAATTTGTAATACCGTTGACATCGAT 435
QY 301 AATATCCTTTTGTGGCTGACGACGTAAATTTGCAACGAATTTCTTGCTAAAGTTCTTCAAT 360
DB 434 AATATCCTTTTGTGGCTGACGACGTAAATTTGCAACGAATTTCTTGCTAAAGTTCTTCAAT 375
QY 361 ATCAACGGCTTAACTATATAATGCTGCAACCGTAAATCAAGCCAGCAACTTTGTGCATTA 420
DB 374 ATCAACGGCTTAACTATATAATGCTGCAACCGTAAATCAAGCCAGCAACTTTGTGCATTA 315
QY 421 CGTATCACTTTTCCGCTGTAATTAATGATAGGTGTAGATTTGTTGCTCAATTTTGGC 480
DB 314 CGTATCACTTTTCCGCTGTAATTAATGATAGGTGTAGATTTGTTGCTCAATTTTGGC 255
QY 481 ACAAAATTTCTAAGCCATTAATTTGACGGCAACATTTAAATCTAATATGATTAATCATAGTA 540
DB 254 ACAAAATTTCTAAGCCATTAATTTGACGGCAACATTTAAATCTAATATGATTAATCATAGTA 195
QY 541 ATGCTAAGCGCTTTATCTAAACCGTCTTGTCCATCATCTCTGTTGTCACATTTGAATTT 600
DB 194 ATGCTAAGCGCTTTATCTAAACCGTCTTGTCCATCATCTCTGTTGTCACATTTGAATTT 135
QY 601 TCCATGTTGAGTTCCTCAATTTCAAGAAATCTTGTCTAAGTTTGTTCATCTCTACTATTAA 660
DB 134 TCCATGTTGAGTTCCTCAATTTCAAGAAATCTTGTCTAAGTTTGTTCATCTCTACTATTAA 75
QY 661 AATTTGGCTCATTTGTACACCTCATATTAAGCACTTTTCTTAATAAGGTAATATATTTTAA 720
DB 74 AATTTGGCTCATTTGTACACCTCATATTAAGCACTTTTCTTAATAAGGTAATATATTTTAA 15
```

```
QY 721 ATTTCGTATTT 731
DB 14 ATANTGNAITTT 4
```

RESULT 5

US-09-134-001C-375/c
Sequence 375, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 375
LENGTH: 669
TYPE: DNA
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-375

Query Match 54.6%; Score 401.8; DB 4; Length 669;
Best Local Similarity 75.0%; Pred. No. 1.2e-91;
Matches 502; Conservative 0; Mismatches 167; Indels 0; Gaps 0;

```
QY 13 TCATCGTATCACATACCCAGCCAGCAACTGTTTCAATCATTTTGTGACGATCGTATGG 72
DB 669 TCATCGTATCACATACCCAGCCAGCAACTGTTTCAATCATTTTGTGACGATCGTATGG 610
QY 73 TTTTAACTTTGTTTCGTAATAATCTTATATAACATCTACGACATTTGTTTCTACTTCACT 132
DB 609 TTTGAGTTTATACGTAATAATACGAATATAACATCAACGACATTCGTTTCTACTTCACT 550
QY 133 ATTATAACCCCATACATGATTTTAAATTTGTTCCCGTTGCATAAATGTTTATTTTTC 192
DB 549 ATTATAACCCCATACATGATTCGAGATTTGTTCCAGTTGACGCTGTTGACGATTTTTC 490
QY 193 AGCTAGAATATAGTAAATCATCTCTGTTTGTGTTTAAATTCATTTCTGCGGCATTTAC 252
DB 489 AGCTAAACATATAGTAAATCGTATTTCTGTTTGTAGTTAAATCTTAAATGATGGCCATTAAC 430
QY 253 CGTCACCTTTAAAGCGTTCTTATCAATTCGTAATACCGTTGACATCGATATATCCTTTTG 312
DB 429 AGTAACCTTTAAAGCGATCTTTATCAATGATTAATCAATGATATCAATCAATCTTTATC 370
QY 313 TGCGTACGACGCTAAATTTGACGCAATCTTGTCTAAAGTTCTTCAATATCAAAACGCTT 372
DB 369 TGCGTCTGTCGCAATACCGCTCTTATTCTTGAAGCAATTTCTTATATCAAAAGGTTT 310
QY 373 AACTATATAATCGTCTGCAACCGTAAATCAAGCCAGCAACTTTGTCTATAGTATCATTTC 432
DB 309 TACAATGTAGTATCTGCGCCCATAGTCCAAATCCAGCTACTTTATCATATATGTCCTCGCTTT 250
QY 433 CGCTGTAATTAATGATAGGTGTAGATTTGTTGTTCTAAATTTTTCGACAAATTTCTAA 492
DB 249 CGCAGTAATGATGATTAATTTGGAGTAGTTGTTTGTGACGAATTTGTCACAAATTTCTAG 190
QY 493 GCCATTAAATTGACGGCAACATTTAAATCTAATATGATTAATCATAGTAATGGCTAAGCGC 552
DB 189 ACCATTATTTTGGAGCAATTAAGTCTAATATATAATAATCATAGGCTTACATAATGC 130
QY 553 TTTATCTAAACCGTCTTGTGCAATCATCTCTGTGTCACATTTGTAATTTTTCATGTGAG 612
DB 129 TTTATCTAAACCCACCTTTTCCATCATTTCTCTATATCAACAGTATAATTTTTCATGAGTAA 70
QY 613 TTCCAATTCAGAAATCTTTGCTAAGTTTGTGTTCTCTACTTCTTAAATTTTTCGCTCAT 672
```


Db 642 CTTTCTTAAGCTTTTGGAGTGTCTACTAGTACGCAAGTCCCAATAAATTCGTA 583
QY 139 ACCCATACATGATTTAAATTTGTCCTGTCATACATGCTTTTATTTTTCAGCTAG 198
Db 582 ATGCCAAACCTCTTAAAGTAACTGTTCTCGATCAATATCTTTTATTTGGAGTTTAGCTAA 523
QY 199 AAGATATAGTAATCATCTCTCTTTTGTAAATTCATTTTTCGCGCAATTTACCGTCAC 258
Db 522 ATAGATGAGTAATTCATCTCTTTAGTGTAGATTACCTTTGATCATCCGCAAGGAC 463
QY 259 TTTAAAGCGTCTTATCAATGTAATACCGTTGACATGATTAATATCTTTTGTGGCTG 318
Db 462 TCGATGAGCGTCAATATCAATCACTAGATGACTAAATTCATTAATATCGCGTCATGCGG 403
QY 319 -----ACGACGTAAATTCGACGAAATCTTCTAA 348
Db 402 TTGCTTTGCTAGTGTGTCAGTGTGTCAGCTTTTAATACGCTTTACTCTCAATAC 343
QY 349 AAGTCTTCTCAATATCAACCGCTTAATATATATATGCTGTGCAACCGTAATCAAGCCGACG 408
Db 342 TACTTCTAGTGTGAGAAAGGCTTAACATATATATATATCTGTCTCCGATTCATCTTTC 283
QY 409 AACTTTGTCAATGATCACTTTTTCGCTGTAATATAATGATAGGTGTAGATTGTTG 468
Db 282 AACACGATTTGTTTCTTACCTTTTGCAGTAAGCATGATGATAGGTGTATCTTTATGTT 223
QY 469 TCTAATTTTCGCAAAATTTCTAAGCCATTAATTTGACGCAACATTAATCAATCAATATGAT 528
Db 222 TCTAAGTTTAGAGCAACTTCAATTCATCCATTTTCCAGGAAGCATTAATCTAAGAGAT 163
QY 529 TAAATCATAGTAATGCTAAGCGCTTTATCTAAACGCTTTGTCATCATCTGTC 588
Db 162 GCAGCATAGTGTGTTTCAAGCAAGCTGCTAGGCTCTTTACCATCTCTGCTTCATG 103
QY 589 CACATTTGTAATTTCAAGTGTAGTGTCCAAATCAAGAAATCTTGTGTAAGTTTGTTCATC 648
Db 102 AATTCAAAAGATCTCTTCTTAATATATATTTAAGTAATCTTCTAATCTATCTTTCATC 43
QY 649 TTTCTACTATTAATAATTT 665
Db 42 ATCAACGATTAATAATTT 26

RESULT 8
US-09-134-000C-415/c
; Sequence 415, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134, 000C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 415
; LENGTH: 555
; TYPE: DNA
; ORGANISM: Enterococcus faecalis
US-09-134-000C-415

Query Match 19.8%; Score 145.8; DB 4; Length 555;
Best Local Similarity 63.0%; Pred. No. 1.3e-27;
Matches 225; Conservative 0; Mismatches 132; Indels 0; Gaps 0;
QY 319 ACAGCTAAATTTGACGAAATTTTGTGTAAGTTCTTCAATATCAAGCGCTTAATCTAT 378
Db 375 ACAGCAAGTAACGCAAGTAAACGAGCTAACAAATTTCTCAATTTGCAATGTTTAAAT 316
QY 379 ATAATCGTCTGCAACCGCTTAATCAAGCCGACCACTTTGTCTATACGTATCACTTTT 438

Db 315 ATAATCATCGCTCCATGCTTAAGCCAGAAACACGGTCAAATTACTGAATCACGCGCAGT 256
QY 439 AATTATATGATAGGTGTAGATTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 498
Db 255 CATCAATAAATTTGGTGTATTTTTCACCTTTGGCGAACACGCGGACATACATCTTAATCCATT 196
QY 499 AATTGACGGCAACATTTAAATCTAATATGATTAATTAATCATAGTAAATGCTAAGCGCTTTATC 558
Db 195 TAAATCTGTTAATCAATCAATCAAGAGATAGATCCCAATTCGTTGTTAAGCGGCTTC 136
QY 559 TAAACCGTCTTGTCCATCATCTCTGTGTCACATTTGTAATTTTTCATGTTGTTGTTGTTGTT 618
Db 135 CAATCTGTACGACCATTTCTGTGTACTTCTGTGTGTATACCCCTCATGTTTAAATTTCAAG 76
QY 619 TTCAAGAAATCTTGTAAATTTTGTTCATCTTCTACTATTAATAATTTGCGTCAATTTG 675
Db 75 CTCAAGAAATCTCGCTAAGTTCTTTTTCATCTTTCAATTAATAATTTGTTGTTGTTGTTGTT 19

Query Match 19.6%; Score 144; DB 4; Length 684;
Best Local Similarity 53.1%; Pred. No. 3.8e-27;
Matches 359; Conservative 0; Mismatches 305; Indels 12; Gaps 2;
QY 6 CGTTTGTCTATCGTATCATATCCCAACGCCAGCACTGTTTCAATCAATTTTGTACAGAT 65
Db 676 CATCTTCTACCTCTAAGATACATCCCATCTCCGGGATGTTTGAATCACTTAGGACTAA 617
QY 66 CGTATGGTTTAACTTGTGTTTAAATATCTTATATAAATCTACGACATTTGTTTCTA 125
Db 616 AATTACTATCTATTTTCTTAGACGTTTAAATAGCGACTTCTACACGTTTGTATCAC 557
QY 126 CTTCACTATTAACCCCATACATGATTTAAATTTGTTCCGTTGTCGTAACAATGTTT 185
Db 556 TATCAAAATTCATATCCCATATTTGAGAAGCAATGAGACGAGGACAGAAATCTCTCCAC 497
QY 186 TATTTTTCAGTAGAGATATAGTAATCATCTCTGTTTGTGTTTGTGTTTGTGTTTGTGTTG 245
Db 496 GTCTAGCATGAATTAATTCATCAATGCGAACTCTTTAGCTGTAAGGTCAATTTCTTTGTC 437
QY 246 CATTTACCGTCACTCTTATCAATTTGTTTAAATCTTATCAATTTGTTTAAATCTGACAT 296
Db 436 CAGCAGCTGTGACGGCGGTTTACGAAGATCAAGTTCAAGATCAGCAATAGTAATATAT 377
QY 297 CGATAATATCTTTTGTGCTGACGAGTAAATTTGCAAGATTTCTGTTAAAGTTCTTT 356
Db 376 TGTATATCTTCTTTTGTGACCGGACGAAAGTAAAGTTTAAATTTGCGGCAAGAAAGTTCTG 317
QY 357 CAATATCAACCGCTTAACTATATATCTCTGACCGTAAATCAACGCCAGCAATTTGT 416
Db 316 CAAAAGCAAAAGCTTTAAGAGATATCAATCACTCTAGTTCCAGGCTTTTAACTCTAT 257
QY 417 CATACGTATCACTTTTGTGCTGTAATTTATATATGATAGGTGTAGATT---GTTGTTGTTAA 473
Db 256 CTTCAATTTGATCAGCTGTGAGAGAAAGGATAGGATAGTTTGTGCGCACTTTTTCGAA 197

RESULT 9
US-09-328-352-4049/c
; Sequence 4049, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 4049
; LENGTH: 684
; TYPE: DNA
; ORGANISM: Acinetobacter baumannii
US-09-328-352-4049

QY 474 TTTTCGCAAAATTTCTAAGCAATTAATTTAGCGGCAACATTAATAATCTAATAATTAAT 533
Db 196 TATCATTAAATGATATTCACCACTCTAATCTCGGTAAACATTACATCTAGTATAATAGAT 137
QY 534 CATAGTAATGCTAAGCGCTTTATCTAAACGCTTTGTCCATCATCTCTGTGTCACAT 593
Db 136 CATATTTCTCAGAAAGAGCTTGATGTTTACCTGATAAGCCATCTGTAAACCCAGCTCTGTA 77
QY 594 TGTAAATTTTCATGTGTAGTCTCAATTCAGAAATCTTGTGCTAAGTTTGTTCATCTCTA 653
Db 76 TATATCTGCTCAGAAAGAGCTTGCTTTAGTATAATCACCAGTTTGTGTTCATCTCA 17
QY 654 CTATTAATAATTTGCGT 669
Db 16 CTAGTAAGATTTCTCAT 1

RESULT 10

US-08-956-171E-822/c
; Sequence 822, Application US/08956171E
; Patent No. 6593114
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; Gil H. Choi
; Patrick S. Dillon
; Craig A. Rosen
; Steven C. Barash
; Michael R. Fannon
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5256
; CORRESPONDENCE ADDRESS:
; ADDRESS: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850

COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/956,171E
; FILING DATE: 20-Oct-1997
; CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 60/009,861
; FILING DATE: January 5, 1996
; APPLICATION NUMBER: 08/781,986
; FILING DATE: January 3, 1997

ATTORNEY/AGENT INFORMATION:

; NAME: Mark J. Hyman
; REGISTRATION NUMBER: 46,789
; REFERENCE/DOCKET NUMBER: PB248P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (240) 314-1224
; TELEFAX: (301) 303-8439

INFORMATION FOR SEQ ID NO: 822:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 760 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 822:

US-08-956-171E-822

Query Match 17.7%; Score 130.2; DB 4; Length 760;

Best Local Similarity 52.6%; Pred. No. 1.1e-23; Indels 35; Gaps 2;
Matches 350; Conservative 0; Mismatches 288;

QY 25 ATACCCAAAGCCGCAAGTCTTCAATCATTTTGTGACGATC-----GTATGGTTTAAAC 79

Db 735 ATACCCAAAGCCGCAAGTCTTGAATCATATGCGCAGCTTCACTAGACACACGATTTAA 676
QY 80 TTGTTTCGTAATAATCTTATATAAACAATCTACGACATTTGTTTCTACTTACTATTATAA 139
Db 675 CTTTTCTTAAGCGTTTAAATAGATTAACAATCTTCTTAAATCACCATAGAAATCAATA 616
QY 140 CCCATACATGATTAATAATTTGTTCCGTTGCAATAATGTTTATTTTATTTTACAGCTAGA 199
Db 615 TGCCAAAGCTCTTTTAAATAATTTTACGGTCAAAATCTTTTATTTTGGTGTGTTTACGCTAAA 556
QY 200 AGATATAGTAATCATACTCTGTTTGTAAATTTCAATTTCTGCGCATTTTACCGTCACT 259
Db 555 TATATTAATAATTTCTGTTCTTTAGGAGTCAAAATTAATCTTCTGATTAATCAGCAAGTACG 496
QY 260 TTAAGCGTCTTATCAATTTGTAATACCGTTTGATCGATAATATCTTTTG----- 312
Db 495 CGATGTGATCATTTATCTATTTCTAAATGTTTAAATTCATCATCATCAGTGGCGTGGT 436
QY 313 -----TGGCTGACGAGCTAAATTTGCAGAAATTTTCGCTAAA 349
Db 435 TCGCTTTGTTCTACAGTTGTAGATTTGCGTTCTTCTTAGAAGTGTCTTAACTCTTAAAGACT 376
QY 350 AGTTCTTCAATATCAAGCGTTAACTATATATCGTCTGACCGTAAATCAAGCCAGCA 409
Db 375 ACTTCTCTTGGTGAATAATGTTTGACGATATAATCATCTGCACAGATTCAAAAACCTTCA 316
QY 410 ACTTTGTATACAGTATCACTTTTTCGTTGTAATTAATATAGTAGGTAGATGTTGTTGT 469
Db 315 ACAGGTTTGTCTTCACTTTAGCAGTCAACATATAATCGGTGTTGTTTATGTTCA 256
QY 470 CTAAATTTTGGCAAAATTTCTAAGCCATTAATTTGACGGCAACATTAATAATTAATTAAT 529
Db 255 CGCAATTTAGTTGCCACCTCGATACCATTCATTTTCAGGCAACATTTAAATCTAGTAGTATG 196
QY 530 AATCATAGTAATGGCTAAGCGTTTATCTAAACCGTCTTGTCCATCATCTCTGTGTCC 589
Db 195 CAAGCATAAATTTCTCCATTTGCAAGTTTCATAAGCTCTTTGGCCATTAATCTGTTCTATGG 136
QY 590 ACATTGTAATTTTCATGTTGAGTTCCAATTTCAAGAAATCTTCTAAGTTTGTGTTCTATCT 649
Db 135 ATTTCAAAAGATTTCTTTCTAAATACATTTTAAAGTAATCTTCTGATTTCTATCTCATCA 76
QY 650 TCTACTATTAATAATTTGCGTC 670
Db 75 TCTACGATAAGTATTTCTGTTIC 55

RESULT 11

US-09-134-000C-2259/c
; Sequence 2259, Application US/09134000C
; Patent No. 6617156

GENERAL INFORMATION:

; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2259
; LENGTH: 696
; TYPE: DNA
; ORGANISM: Enterococcus faecalis
; US-09-134-000C-2259

Query Match 17.5%; Score 129; DB 4; Length 696;

Best Local Similarity 54.0%; Pred. No. 2.2e-23; Indels 21; Gaps 1;
Matches 301; Conservative 0; Mismatches 235;

FILE REFERENCE: 032796-032
CURRENT APPLICATION NUMBER: US/09/134,000C
PRIOR FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/055,778
PRIOR FILING DATE: 1997-08-15
NUMBER OF SEQ ID NOS: 6812
SOFTWARE: PatentIn version 3.1
SEQ ID NO 1740
LENGTH: 696
TYPE: DNA
ORGANISM: Enterococcus faecalis
US-09-134-000C-1740

Query Match 17.1%; Score 125.8; DB 4; Length 696;
Best Local Similarity 53.8%; Pred. No. 1.4e-22;
Matches 365; Conservative 0; Mismatches 287; Indels 27; Gaps 4;

QY 20 ATCCATACCCAGCCGACGACTGTTCAATCATTTTGTCCAGCATCGTATGTTT--- 76
DB 683 ATCTTATATCCAACTCCCCACACTGTTGAATGACTTTTTCACCACTGTTCTTCA 624
QY 77 AACCTGTTTCGTAATATCTTATAAACAATCTACGACATTTGTTTCTACTTCACTATTA 136
DB 623 ATCTTATCTTAAATGACTGAGTGAACCATACGTTTATAGCAGAACATCTTCT 564
QY 137 TAACCCCATACATGATTTAAATTTGTTCCCGTTGCATAACATGTTTATTTTCAGCT 196
DB 563 TGTTCACACACAGCTCAAAAATTTTATCCGCACTAAACACACAGATTTGGATGACTAGCT 504
QY 197 AGAAGATATAGTAATCATCTCTGTTTGTATTTTATTTTCAATTTTTCGCGCATTTACCG-- 254
DB 503 AATAGATAGGATACCAAACTCAAGCGGTTAATTTGAATTTCTTTTCTTCTATTGTT 444
QY 255 -TCACCTTTAAAGCGTTCTTATCAATTTGAATACCGTTTGACATCGATATATCTCTTTT-- 311
DB 443 TTGACTTCTGAGAGCTCTTATTAATCATTAATGCGCGACTTCTAATTCATCGGGTCT 384
QY 312 -----GTGGCTGACGACGTAATAATTTGCAAGAAATTTGCAAGAAATTTGCTAAAGTTCT 355
DB 383 TCTTTGGTTAATTGCAATTTGGCTTCTACGTAATAATAGACTTTTACCGAGCCATTACTTCT 324
QY 356 TCAATATCAACCGCTTAATCATATATATCGTTCACCGTAATCAAGCCAGCACTTTG 415
DB 323 AAAGGATTAAGAGGCTTGGTGAATAATCGTTCACCGAGCACTTAAGCCTTTGATTTTA 264
QY 416 TCATACGATATCACTTTTCGCTGTAATATAATGATAGGTGTAGATTGTTGTTGCTAATT 475
DB 263 TCCATATCGGTTGTTTTCGCTGTAATATAATTAATTTGAAATTTGCGATTCTTTTCCGAAGT 204
QY 476 TTGCGACAAATTTCTAAGCCATTAAATGACGGCAACATTAAATCTAATATGATTAATC- 534
DB 203 TCTTTCACTACTTCCATTCATCCATGATTTGGCATCATAAATGTCGAAGATTAAATACG 144
QY 535 --ATAGTAATGGCTAAGCGCTTTATCTAAACCGCTTGTCCATCATCTCTGTGTCCACA 592
DB 143 ATATCTGGTGTGTGTGCAATTTGGATAACCGCTTCTTTTCCATCAATAAGCTTTTACGACT 84
QY 593 TTGTAATTTTCATGTGTAGTTTCCAAATTCAGAAATCTTGTAAAGTTTGTTCATCTTCT 652
DB 83 TCGTATCTCTTATATGTAATATAAATCTCAATAGTTTCTTACAATTTCTTTATCATCATCT 24
QY 653 ACTATTAATAATTTGGGTCA 671
DB 23 GCAACTAAAATTTTTCATAA 5

RESULT 14

US-09-082-077-2/c
Sequence 2, Application US/09082077
Patent No. 6514746
GENERAL INFORMATION:
APPLICANT: Microcide Pharmaceuticals, Inc.
TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS HISTIDINE PROTEIN KINASE

TITLE OF INVENTION: ESSENTIAL GENES
FILE REFERENCE: SR Silverstein Microcide 234/067US
CURRENT APPLICATION NUMBER: US/09/082,077
CURRENT FILING DATE: 1998-05-20
PRIOR APPLICATION NUMBER: US 60/009,102
PRIOR FILING DATE: 1995-12-22
PRIOR APPLICATION NUMBER: PCT/US97/23912
PRIOR FILING DATE: 1997-12-23
PRIOR APPLICATION NUMBER: US 08/713,718
PRIOR FILING DATE: 1996-09-13
NUMBER OF SEQ ID NOS: 5
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 702
TYPE: DNA
ORGANISM: Staphylococcus aureus
US-09-082-077-2

Query Match 17.1%; Score 125.6; DB 4; Length 702;
Best Local Similarity 52.3%; Pred. No. 1.6e-22;
Matches 324; Conservative 0; Mismatches 269; Indels 27; Gaps 1;

QY 73 TTTTAACTGTTTCTGTAATATCTTATATAAATCTTACGACATTTGTTTCTACTTCACT 132
DB 630 TTCAATCTTTTTCACGTAACGACGAAATCGTTTACATCGACCGTAGCATCGCCAAAGTA 571
QY 133 ATTATAACCCCATACATGATTTAAATTTGTTCCGTTGTCATAACATCGTTTATTTTTC 192
DB 570 ATCATAGCCCATACTGTTTGTATTAATGTTACGTTGTCATTACTTGTCCCATATGTTT 511
QY 193 AGCTAGAAGATATAGTAATCATCTCTGTTTGTAAATTTCAATTTCTGCGCCATTTAC 252
DB 510 TGATAAATAATGGAACAATTCAAATTCAGATGTTGTTAATTCATATCTTCGCCACGTTT 451
QY 253 CGTCACTTTAAAGGTTCTTATCAATTTGTAATACCGTTTGACATCGATAATATCTTTTG 312
DB 450 TTTAATAGAATATCGTCTGGATAAATCAAAATATCTTTAATTTGATTTTCATTCGTTTAC 391
QY 313 TGGC-----TGACGAGCTAAATTTGCACGAATTTCTTTCG 345
DB 390 ATTTCCAGTGTCTTGTGCTGTTGTGAGTAAATGACGAGTAAAGTTTCGTTTTCACAGTGC 331
QY 346 TAAAGTTTCTTCAATATCAAAACGGCTTAACTATATATAATCGTTCGACCGTAATCAAGCCC 405
DB 330 GATTAATTCACGCTACTATAACGGTTTCGTTTACATAGTCATCTGCACCTAGTTCTTAACC 271
QY 406 AGCACTTTGTCTACGATATCATCTTTTCGCTGTAATTAATATGATAGGTGTAGATTGTTG 465
DB 270 AAGCACTTTATCAATTTCTGAATCTTTAGCAGTAAGCAATTTATTTGTCATTTCTGTAATT 211
QY 466 TTGCTAATTTTTCGACAAATTTCTAAGCCATTAATTTGACGGCAACATTAATATCTAATAT 525
DB 210 TTTGCGCACTTCAGACATCTTCCATACCATACGACGAGGTAAATATCTAGTAA 151
QY 526 GATTAATATCATAGTAATGAGCTTAAGCGCTTTATCTTAAACCGCTTTGTCCATCATCTCTGT 585
DB 150 TACGATGCTGTTCTTCTTCAATTAATTAAGTCGATCATTTACCATCGTATGCACA 91
QY 586 GTCCACATTTGAATTTTTCATGTTGAGTTTCAATTTCAAGAAATCTTGTCAAGTTTGTTC 645
DB 90 GTACACATCGTATCTCTTTTAAAGTTAAATTTCTAAAAATATATCAGCAATCGGTTTTTC 31
QY 646 ATCTTCTACTATTAAATTT 665
DB 30 ATCATCAACTACAACTT 11

RESULT 15

US-09-082-077-1/c
Sequence 1, Application US/09082077
Patent No. 6514746
GENERAL INFORMATION:
APPLICANT: Microcide Pharmaceuticals, Inc.

Search completed: October 5, 2004, 09:25:37
Job time : 61.6323 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 4, 2004, 17:59:28 ; Search time 99.6239 Seconds
(without alignments)
1428.359 Million cell updates/sec

Title: US-09-006-627-2
Perfect score: 2291
Sequence: 1 MTKKLRNNWIIVTMTIFV.....CGSIKIKSEINKGTFPKIIF 451

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 25.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phase.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertibrate.*
- 14: sp_unclassified.*
- 15: sp_virus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2291	100.0	451	16	Q9KJN3
2	1640.5	71.6	456	16	Q8CSL7
3	710.5	31.0	483	16	Q92BX7
4	703.5	30.7	483	16	Q9RPF9
5	621.5	27.1	502	16	Q83Q7
6	606.5	26.5	491	16	Q9CF88
7	509.5	25.7	490	2	O07382
8	563.5	24.6	547	16	Q88WS3
9	547	23.9	462	16	Q8R809
10	533.5	23.3	500	2	Q9LAU2
11	532.5	23.2	500	16	O87528
12	532.5	23.2	500	16	Q8P2J7
13	532.5	23.2	500	16	Q87BB3
14	523.5	22.9	500	16	Q8K8K0
15	516	22.5	501	16	Q8E3T5
16	515	22.5	501	16	Q8DY69

17	493	21.5	452	16	Q81R45
18	492.5	21.0	467	16	Q8XHZ3
19	481.5	21.0	459	16	Q9KFU8
20	464.5	20.3	482	16	Q97FH8
21	463.5	20.2	463	16	Q8ESN1
22	452.5	19.8	441	16	Q9CGT8
23	452.5	19.8	472	16	Q8YXW0
24	445	19.4	484	16	Q81I37
25	437.5	19.1	441	2	O07386
26	437.5	19.1	487	16	Q81VB3
27	435	19.0	463	16	Q81HX8
28	429	18.7	613	16	Q814H3
29	424	18.5	613	16	Q81J31
30	423.5	18.5	583	2	Q9L523
31	423.5	18.5	588	16	Q8NWF3
32	422.5	18.4	588	16	Q99TZ9
33	416	18.2	463	16	Q81V50
34	414.5	18.1	262	2	Q9EXE5
35	413.5	18.0	591	16	Q92723
36	412.5	18.0	609	16	Q836C1
37	412.5	18.0	611	2	Q9REA6
38	411.5	18.0	458	16	Q88ZM9
39	411	17.9	624	16	Q890H9
40	410.5	17.9	577	16	Q8R6U6
41	409.5	17.9	405	16	Q894Q4
42	409	17.9	469	16	Q8R755
43	407.5	17.8	584	16	Q8EQ91
44	405.5	17.7	566	16	Q9RWC5
45	404.5	17.7	468	16	Q89916

ALIGNMENTS

RESULT 1

Q9KJN3	ID	Q9KJN3	PRELIMINARY;	PRT;	451 AA.
AC	Q9KJN3;				
DT	01-OCT-2000	(TrEMBLrel. 15, Created)			
DT	01-OCT-2000	(TrEMBLrel. 15, Last sequence update)			
DT	01-OCT-2003	(TrEMBLrel. 25, Last annotation update)			
DE	Hypothetical protein arls.				
GN	ARLS OR SAV1414 OR SAI246 OR MW1304.				
OS	Staphylococcus aureus (strain Mu50 / ATCC 700699),				
OS	Staphylococcus aureus (strain N315),				
OS	Staphylococcus aureus, and				
OS	Staphylococcus aureus (strain MW2).				
OC	Bacteria; Firmicutes; Bacillales; Staphylococcus.				
OX	NCBI_TaxID=158878, 158879, 1280, 196620;				
[1]	SEQUENCE FROM N.A.				
RP	STRAIN=ISP794;				
RC	MEDLINE=20327575; PubMed=10869073;				
RA	Fournier B., Hooper D.C.;				
RT	"A new two-component regulatory system involved in adhesion,				
RT	autolysis, and extracellular proteolytic activity of Staphylococcus				
RT	aureus.";				
RL	J. Bacteriol. 182:3955-3964(2000).				
[2]	SEQUENCE FROM N.A.				
RP	STRAIN=Mu50, and N315;				
RC	MEDLINE=21311952; PubMed=11418146;				
RA	Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,				
RA	Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,				
RA	Kanamori M., Matsunaru H., Maruyama A., Murakami H., Hosoyama A.,				
RA	Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,				
RA	Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,				
RA	Kanehisa M., Yanashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,				
RA	Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;				
RT	"Whole genome sequencing of methicillin-resistant Staphylococcus				
RT	aureus.";				
RL	Lancet 357:1225-1240(2001).				
RL	[3]				

Q81r45 bacillus an
Q8xhz3 clostridium
Q9kfu8 bacillus ha
Q97fh8 clostridium
Q8esn1 oceanobacil
Q9cgt8 lactococcus
Q8yxw0 anabaena sp
Q81i37 bacillus ce
O07386 lactococcus
Q81vb3 bacillus an
Q81hx8 bacillus ce
Q814h3 bacillus ce
Q81j31 bacillus an
Q9l523 staphylococ
Q8nwf3 staphylococ
Q99tz9 staphylococ
Q81v50 bacillus an
Q9exes lactobacill
Q92723 listeria in
Q836c1 enterococcu
Q9rea6 enterococcu
Q88zm9 lactobacill
Q890h9 lactobacill
Q8r6u6 thermoanaer
Q894q4 clostridium
Q8r755 thermoanaer
Q8eq91 oceanobacil
Q9rwc5 deinococcus
Q89916 clostridium


```
RP SEQUENCE FROM N.A.
RC STRAIN=WW2;
RX MEDLINE=22040717; PubMed=12044378;
RA Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,
RA Nagai Y., Iwama N., Asano K., Naimi T., Kuroda H., Cui L.,
RA Yamamoto K., Hiramatsu K.;
RT "Genome and virulence determinants of high virulence community-
RT acquired MRSA.";
RL Lancet 359:1819-1827(2002).
CC -1- SIMILARITY: TO PROKARYOTE SENSORY TRANSDUCTION PROTEINS.
DR ENBL; AF165314; AAF85897.1; -.
DR ENBL; AP003362; BAB57576.1; -.
DR ENBL; AP003133; BAB42506.1; -.
DR ENBL; AP004826; BAB95169.1; -.
DR PIR; F89918; F89918.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0000155; F:two-component sensor molecule activity; IEA.
DR GO; GO:0007600; F:sensory perception; IEA.
DR GO; GO:0007165; P:signal transduction; IEA.
DR InterPro; IPR003594; ATPbind_ATPase.
DR InterPro; IPR004358; Bact_sens_pr_C.
DR InterPro; IPR003660; HAMP_sens_pr_C.
DR InterPro; IPR005467; His_kinase.
DR InterPro; IPR003661; His_kinase.
DR Pfam; PF00672; HAMP; 1.
DR Pfam; PF02518; HATPase_c; 1.
DR Pfam; PF00512; Hiska; 1.
DR PRINTS; PR00344; BCTRLSENSOR.
DR SMART; SM00304; HAMP; 1.
DR SMART; SM00387; HATPase_c; 1.
DR SMART; SM00388; Hiska; 1.
DR PROSITE; PS00109; HIS_KIN; 1.
DR PROSITE; PS0109; HIS_KIN; 1.
KW Hypothetical protein; Kinase; Phosphorylation; Sensory transduction;
KW Transferase; Complete proteome.
SQ SEQUENCE 451 AA; 52400 MW; 6308576067A22438 CRC64;

Query Match 100.0%; Score 2291; DB 16; Length 451;
Best Local Similarity 100.0%; Pred. No. 1.2e-113;
Matches 451; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTKRKLNNWIIIVTTMTFTVFIFLFCIIIPFLKDTLHNSLDDAERSDDINLPHSKP 60
DB 1 MTKRKLNNWIIIVTTMTFTVFIFLFCIIIPFLKDTLHNSLDDAERSDDINLPHSKP 60

QY 61 VKDISALDLNASLGNFOEIIYDEHNKLPFSDNTVVRPEGVYHRYFDRVTKRYKG 120
DB 61 VKDISALDLNASLGNFOEIIYDEHNKLPFSDNTVVRPEGVYHRYFDRVTKRYKG 120

QY 121 EYLIKEPITTDQPKGYSLHLSLENDYDNVKSLYIALAFGLVIATITATISYVFSTQ 180
DB 121 EYLIKEPITTDQPKGYSLHLSLENDYDNVKSLYIALAFGLVIATITATISYVFSTQ 180

QY 181 TKPLVSLNKMIEIRRDGFQNKQLQNTNVEEDNLANTFNEMSQIESFNQORQFVEDA 240
DB 181 TKPLVSLNKMIEIRRDGFQNKQLQNTNVEEDNLANTFNEMSQIESFNQORQFVEDA 240

QY 241 SHELRTPLOIIQGHNLNLIQRWKKDPVAVLESNLNIEEMNRRIIKLVEELLETKGDVND 300
DB 241 SHELRTPLOIIQGHNLNLIQRWKKDPVAVLESNLNIEEMNRRIIKLVEELLETKGDVND 300

QY 301 ISSEAQTVDHINDEIRSRHSLKQLHPDYQFDTDLTLSKNLEIKMKPHQFQFLIFIDNAI 360
DB 301 ISSEAQTVDHINDEIRSRHSLKQLHPDYQFDTDLTLSKNLEIKMKPHQFQFLIFIDNAI 360

QY 361 KYDVKNKKIKVTRLKNKQKIIEITDHGIGIPEDQDFIFDRFYRVDKSRSRSGGNGLG 420
DB 361 KYDVKNKKIKVTRLKNKQKIIEITDHGIGIPEDQDFIFDRFYRVDKSRSRSGGNGLG 420

QY 421 LSIAQKIIQLNGSGSIKIKSEINKGTTFKIIIF 451

Db 421 LSIAQKIIQLNGSGSIKIKSEINKGTTFKIIIF 451

RESULT 2
Q8CSL7 PRELIMINARY; PRT; 456 AA.
ID Q8CSL7
AC Q8CSL7, 2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN SE1099.
OS Staphylococcus epidermidis.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1282;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 12228;
RA Zhang Y., Ren S., Li H., Fu G., Lu L., Lu G., Jia J., Tu Y., Qin Z.,
RA Chen Z., Wen Y.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
DR ENBL; AB016747; AA04696.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0007165; P:two-component sensor molecule activity; IEA.
DR GO; GO:0007600; P:sensory perception; IEA.
DR InterPro; IPR003594; ATPbind_ATPase.
DR InterPro; IPR004358; Bact_sens_pr_C.
DR InterPro; IPR003660; HAMP_sens_pr_C.
DR InterPro; IPR005467; His_kinase.
DR InterPro; IPR003661; His_kinase.
DR Pfam; PF00672; HAMP; 1.
DR Pfam; PF02518; HATPase_c; 1.
DR PRINTS; PR00344; BCTRLSENSOR.
DR SMART; SM00387; HATPase_c; 1.
DR SMART; SM00388; Hiska; 1.
DR PROSITE; PS00109; HIS_KIN; 1.
DR PROSITE; PS0109; HIS_KIN; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 456 AA; 52859 MW; 8ACF8A6CF9D7F77B CRC64;

Query Match 71.8%; Score 1640.5; DB 16; Length 456;
Best Local Similarity 69.9%; Pred. No. 3.1e-79;
Matches 316; Conservative 68; Mismatches 67; Indels 1; Gaps 1;

QY 1 MTKR-KLRNNWIIIVTTMTFTVFIFLFCIIIPFLKDTLHNSLDDAERSDDINLPHSK 59
DB 1 MKRQKLKYNWIIIVTTMTFTVFIFLFCIIIPFLKDTLHNSLDDAERSDDINLPHSK 60

QY 60 PVKDISALDLNASLGNFOEIIYDEHNKLPFSDNTVVRPEGVYHRYFDRVTKRYKG 119
DB 61 SLSDISALDLNASLGNFOEIIYDVKGRKLIQTSNDNTLAYDNKIDFKKPERHIQRSHG 120

QY 120 IEYLIKEPITTDQPKGYSLHLSLENDYDNVKSLYIALAFGLVIATITATISYVFSTQ 179
DB 121 INYLVTEIRSDKFGSYGLVHSLQNDLVKSLYIALAFGLVIATITATISYVFSSQ 180

QY 180 TKPLVSLNKMIEIRRDGFQNKQLQNTNVEEDNLANTFNEMSQIESFNQORQFVED 239
DB 181 TKPVTNKNKNQIIRRDGFQNKLETTTYEEDNLIIDFNEMMYQIESFNQORQFVED 240

QY 240 ASHELRTPLQIIQGHNLNLIQRWKKDPVAVLESNLNIEEMNRRIIKLVEELLETKGDVN 299
DB 241 ASHELRTPLQIIQGHNLNLIQRWKKDPVAVLESNLNIEEMNRRIIKLVEELLETKDRVN 300

QY 300 DISSEAQTVDHINDEIRSRHSLKQLHPDYQFDTDLTLSKNLEIKMKPHQFQFLIFIDNA 359
DB 301 HNVLECEVDVNSEIQSRVKSQHLHPDYTFETHLATKPIQLKINRHOEQQLLIIFDNA 360

QY 360 IKYDVKNKKIKVTRLKNKQKIIEITDHGIGIPEDQDFIFDRFYRVDKSRSRSGGNGL 419
```

```

Db 361 MYDTEHGHKIKVTQKNNIMIDTDHGMGPKADLEFIDRFYRVKRSRSGGNGL 420
Qy 420 GLSIAQKIQIQLNGSGIKKSEKNGTFFKIIF 451
Db 421 GLSIAEKIVQLNGMGMQVESELQKTYTTFKISF 452

RESULT 3
Q2BX7
ID Q92BX7 PRELIMINARY; PRT; 483 AA.
AC Q92BX7;
DT 01-DEC-2001 (TRENBLrel. 19, Created)
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Two-component sensor histidine kinase.
GN LISK OR LIN1415.
OS Listeria innocua.
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OX NCBI_TaxID=1642;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CLIP 11262 / Serovar 6a;
RX MEDLINE=21537279; PubMed=11679669;
RA Glaeser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
RA Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA Jones L.-M., Kaerst U., Kref J., Kuhn M., Kunst F., Kurapkat G.,
RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
RA Nordsiek G., Novella S., de Pablo B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schluster T., Simoes N., Tierrez A.,
RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.,
RT "Comparative genomics of Listeria species.";
RL Science 294:849-852(2001).
DR EMBL; AL596169; CAC96646.1; -.
DR PIR; AF1609; AF1609.
DR ListList; LIN01415; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; P:ATP binding; IEA.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0000155; F:two-component sensor molecule activity; IEA.
DR GO; GO:0007165; P:signal transduction; IEA.
DR InterPro; IPR003594; ATPbind_ATPase.
DR InterPro; IPR004358; Bact_sens_pr_C.
DR InterPro; IPR003660; HAMK.
DR InterPro; IPR005467; His_kinase.
DR InterPro; IPR003661; His_kinA_N.
DR Pfam; PF00672; HAMK; 1.
DR Pfam; PF02518; HAMK; 1.
DR Pfam; PF00512; HisKA; 1.
DR PRINTS; PR00344; BCTRLSENSOR.
DR SMART; SM00304; HAMK; 1.
DR SMART; SM00387; HAMK; 1.
DR SMART; SM00388; HisKA; 1.
DR PROSITE; PS50885; HAMK; 1.
DR PROSITE; PS50109; HIS_KIN; 1.
KW Kinase; Complete proteome.
SQ SEQUENCE 483 AA; 55501 MW; F5712DF5304C32D2 CRC64;

Query Match
Best Local Similarity 31.0%; Score 710.5; DB 16; Length 483;
Matches 163; Conservative 102; Mismatches 169; Indels 47; Gaps 10;

Qy 1 MTKKRLNNWIVTMTIFLFLCLII-----IFFLKDTLHNS 41
Db 8 LKRSRLKPKWFGASAAIFLTFLLFSYALYQIGQMLNBEPEVKELLATTSTLTNQD 67
Qy 42 LDDAERSDDINLPHSPKVDKISALDNASIGNPQEIIDHN---NKLFTSNDNTV 98
Db 68 LTDNE-----EIKYLFNNDKT-----VNRKLQD-QVINLYDKDGHFINKYFRNQDIT 115

```

RESULT 4

Q9RPY9

ID Q9RPY9

AC Q9RPY9

DT 01-MAY-2000

DT 01-MAY-2000

DT 01-OCT-2003

DE Lisk

GN Lisk

OS Listeria

OC Bacteria

OX NCBI_TaxID=1639;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=LO28 / Serovar 1/2c;

RX MEDLINE=20011355;

RA Cotter P.D., Emerson N., Gahan C.G., Hill C.;

RT "Identification and disruption of lisk, a genetic locus encoding a

RT two-component signal transduction system involved in stress tolerance

RL and virulence in Listeria monocytogenes.";

J. Bacteriol. 181:6840-6843(1999).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=EGD-e / Serovar 1/2a;

RX MEDLINE=21537279; PubMed=11679669;

RA Glaeser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,

RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,

RA Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,

RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,

RA Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,

RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,

RA Jones L.-M., Kaerst U., Kref J., Kuhn M., Kunst F., Kurapkat G.,

RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,

RA Nordsiek G., Novella S., de Pablo B., Perez-Diaz J.-C., Purcell R.,

RA Remmel B., Rose M., Schluster T., Simoes N., Tierrez A.,

RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;

RT "Comparative genomics of Listeria species.";

RL Science 294:849-852(2001).

CC -1- SIMILARITY: TO OTHER PROKARYOTIC SENSORY TRANSDUCTION HISTIDINE

KINASES.

DR EMBL; AF139908; AAF03933.1; -.

DR EMBL; AL591978; CAC99456.1; -.

DR PIR; AB1247; AB1247.

```
DR ListList; LMO01378; -.
DR GO: 0016020; C:membrane; IEA.
DR GO: 0005524; F:ATP binding; IEA.
DR GO: 0016301; F:kinase activity; IEA.
DR GO: 0016740; F:transferase activity; IEA.
DR GO: 0000155; F:two-component sensor molecule activity; IEA.
DR GO: 0007600; P:sensory perception; IEA.
DR GO: 0007165; P:signal transduction; IEA.
DR InterPro: IPR003594; ATPbind_ATPase.
DR InterPro: IPR004358; Bact_sens_pr_C.
DR InterPro: IPR003660; HAMP.
DR InterPro: IPR005467; His_kinase.
DR InterPro: IPR003661; His_kinase_N.
DR Pfam: PF00518; HAMP; 1.
DR Pfam: PF02518; HAMP; 1.
DR PRINTS; PR00344; BCTRISENSOR.
DR SMART; SM00304; HAMP; 1.
DR SMART; SM00388; HSKA; 1.
DR PROSITE; PS00885; HAMP; 1.
DR PROSITE; PS01019; HIS_KIN; 1.
KW Kinase; Phosphorylation; Sensory transduction; Transferase;
KW Complete proteome.
SQ SEQUENCE 483 AA; 55460 MW; B477C2C3D653DC95 CRC64;

Query Match 30.78; Score 703.5; DB 16; Length 483;
Best Local Similarity 33.78; Pred. No. 1.2e-29;
Matches 162; Conservative 102; Mismatches 170; Indels 47; Gaps 10;

QY 1 MTKRKLNNIIIVTMTITFTIFLFLCLII-----IPFLKDTLHNSR 41
DB 8 LKRSRLFKFTFGASAIITFTFLFSYAIQGIQGLMLNEEPEVKELLATSTTNQD 67
QY 42 LDDAERSSDINNLFHSPKVKDISALDNASLGNFBIILYDEHN---NKLFTSDNTV 98
DB 68 LTDE-----EIKYLFNDKT-----VNRKLOD-QVINLYDKDGHFINKYFERSQDIT 115
QY 99 RVFPGYEHRYF-----DRVI--KRYKGYEYLIIKEPIITQD---FKGYSLLIHSLNVDN 149
DB 116 SID---FSQYFVSGTDFKFNKPTIDGQKMTAQMPIVADDNTTVIGYQAVNPLTSYNR 172
QY 150 IVKSLYIIALAFGVIAITITATISYVFSTQITKPLVSLSNKMTIEIRRDGFONKQLNTN- 208
DB 173 MDRLLVMTLLGAVALFISGMLGYLLAQNFLPLRLATMDIRKNGFKRIETKNS 232
QY 209 YEEIDNLANTFNMMSQIEESFNQORQFVEDASHELRTPLQIIQGHNLNLIQWKKDPAY 268
DB 233 RDEIGELTVVNDMTRIETSPQOKQFVEDASHELRTPLQIMEGHLKLLTRNGKDDPAV 292
QY 269 LBSNLISIEEMRIIKLVELELTGKDVNDISSEAOVTHINDEIRSRHSLKQLHPDY 328
DB 293 LDESNAASLELMRKVLQVEMLDLSRAEQISQTKELQITDVNATVEQVRNFMVYENF 352
QY 329 QFDTDLTSKNLEIKMKPHQEQFLIFIDNAIKYDVGNKKIKVKTRLNKKIIEITDHG 388
DB 353 TFLKEDDTDLRALIQNHLEQIIILIMDNVAKYSGDGTVDHMYKVEQKQIHDVVDYG 412
QY 389 IGIPEDQDFIDRFYRVDSRSRQSGNGGLGSLAQKIIQLNGGSIKIKSEINKGTTFK 448
DB 413 EGISQEBIDKIFNFRYRVDSKRSREKGGNGGLGLAIQKLVGVLGTLINAVSEPKGTITK 472
QY 449 I 449
DB 473 I 473

RESULT 5
Q836Q7 PRELIMINARY; PRT; 502 AA.
AC Q836Q7
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Sensor histidine kinase.
GN EF1051.
OS Enterococcus faecalis (Streptococcus faecalis).
OC Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.
OX NCBI_TaxID=1351;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=V583 / ATCC 700802;
RX MEDLINE=22550857; PubMed=12663927;
RA Paulsen I.T., Banerjee L., Myers G.S.A., Nelson K.E., Seshadri R.,
RA Read T.D., Fouts D.E., Eisen J.A., Gill S.R., Heidelberg J.F.,
RA Tettelin H., Dodson R.J., Umayam L., Brinkac L., Beanan M.,
RA Daugherty S., DeBoy R.T., Durkin S., Kolonay J., Madupu R., Nelson W.,
RA Vamathevan J., Tran B., Upton J., Hansen T., Shetty J., Khouri H.,
RA Utterback T., Radune D., Ketchum K.A., Dougherty B.A., Fraser C.M.;
RT "Role of mobile DNA in the evolution of vancomycin-resistant
RT Enterococcus faecalis."
RL Science 299:2071-2074(2003).
DR EMBL; AE016950; AAO80854.1; -.
DR TIGR; EF1051; -.
DR GO: 0016020; C:membrane; IEA.
DR GO: 0005524; F:ATP binding; IEA.
DR GO: 0016301; F:kinase activity; IEA.
DR GO: 0000155; F:two-component sensor molecule activity; IEA.
DR GO: 0007165; P:signal transduction; IEA.
DR InterPro: IPR003594; ATPbind_ATPase.
DR InterPro: IPR004358; Bact_sens_pr_C.
DR InterPro: IPR003660; HAMP.
DR InterPro: IPR005467; His_kinase.
DR InterPro: IPR003661; His_kinase_N.
DR Pfam: PF00672; HAMP; 1.
DR Pfam: PF02518; HAMP; 1.
DR PRINTS; PR00344; BCTRISENSOR.
DR SMART; SM00304; HAMP; 1.
DR SMART; SM00388; HSKA; 1.
DR PROSITE; PS00885; HAMP; 1.
DR PROSITE; PS01019; HIS_KIN; 1.
KW Kinase; Complete proteome.
SQ SEQUENCE 502 AA; 58308 MW; A922E5C78FE04D74 CRC64;

Query Match 27.18; Score 621.5; DB 16; Length 502;
Best Local Similarity 30.08; Pred. No. 2.8e-25;
Matches 147; Conservative 110; Mismatches 164; Indels 69; Gaps 12;

QY 10 WIIVTMTITFTIFLFLCLIIIPFLKDTLHNSLDDAERSSDINNLF---HSPKVKDI- 64
DB 18 WAFASFFIFVVTIPAVITYKSSVSLIVAKENVEATIAEVTNLANANENLTVTDVF 77
QY 65 -----SALDLN-----ASLGNFQ-----IIYDEHNKLFSTSDNTV 98
DB 78 DYLLKTPSERDENYNNKHTAVEGSMEMDSFISLQGPFLYLSVYDTNQKLVFKTQNE--- 134
QY 99 RVFPGYEHRYFDRVI-----KRYKGYEYLIIKEPIITQDQKGYSLLIHSL 144
DB 135 -----YDKLLQDLQPLPVRTVFDKTFGYSVEPIFSKE---TREKIGYIQAPEL 181
QY 145 ENYDNIVKSLYIIALAFGVIAITITATISYVFSTQITKPLVSLSNKMTIEIRRDGFONKL- 203
DB 182 SSFYERNHLLTLVLEVLISLVSSVLGLFLLSKPLKVLKVLRTMDTIRKDP-QSDVH 240
QY 204 --QLNTNYEIDNLANTFNMMSQIEESFNQORQFVEDASHELRTPLQIIQGHNLNLIQW 261
DB 241 MPEINTR-DELADISEIFNEMLMRMYIEQQEQFVEDVSHELRTPLVAMEGHLNLRW 299
QY 262 GKQDPAVLEESLNISIEEMRIIKLVELELTGKDVNDISSEAOVTHINDEIRSRHSL 321
DB 300 GKDDPEILDES LKASLQIESRMKSLVOEMLDLSRAEQVDTQVANERTDAKVYQVFNFF 359
QY 322 KQLHPDYQ--FDTDLTSKNLEIKMKPHQEQFLIFIDNAIKYDVGNKKIKVKTRLNKKQ 379
```

```

Db 360 QLVYEFHITLDDDLPTPE-VELKIYRNHFEQLLIILLNNAIKYSTDRKEVHISISRTWNE 418
Qy 380 KIIETDHGIGIPEDQDFIDRFYRVKDSRSRSGNGGLGSLIAQKIIQLNGSGIKIKS 439
Db 419 FEIAVDQFGEGITEDEKIFDRFYRVKDSRSRSGNGGLGSLIAQKIIQLNGSGIKIKS 478
Qy 440 EINKGTTTFKI 449
Db 479 VLHQGTIFRI 488

RESULT 6
Q9CF88 PRELIMINARY; PRT; 491 AA.
AC Q9CF88;
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Sensor protein kinase.
GN KINA OR L11593.
OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
OX NCBI_TaxID=1360;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=IL1403;
RX MEDLINE=21235186; PubMed=1137471;
RA Bolotin A., Wincker P., Mauger S., Jaillon O., Malarne K.,
RA Weissenbach J., Ehrlich S.D., Sorokin A.;
RT "The complete genome sequence of the lactic acid bacterium Lactococcus
RT lactis ssp. lactis IL1403."
RL Genome Res. 11:731-753(2001).
CC -!- SIMILARITY: TO PROKARYOTE SENSORY TRANSDUCTION PROTEINS.
DR EMBL; AE006390; AAK05691.1; -.
DR PIR; A86824; A86824.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0000155; F:two-component sensor molecule activity; IEA.
DR GO; GO:0007165; P:sensory perception; IEA.
DR InterPro; IPR003594; ATPbind_ATPase.
DR InterPro; IPR004358; Bact_sens_pr_C.
DR InterPro; IPR003660; HAMP.
DR InterPro; IPR005467; His_kinase.
DR InterPro; IPR003661; His_kinA_N.
DR Pfam; PF00672; HAMP; 1.
DR Pfam; PF02518; HATPase_C; 1.
DR PRINTS; PR00512; HisKA; 1.
DR PRINTS; PR00344; BCTRLSENSOR.
DR SMART; SM00304; HAMP; 1.
DR SMART; SM00387; HATPase_C; 1.
DR SMART; SM00388; HisKA; 1.
DR PROSITE; PS50885; HAMP; 1.
DR PROSITE; PS50109; HIS_KIN; 1.
KW Kinase; Phosphorylation; Sensory transduction; Transferase;
KW Complete proteome.
SQ SEQUENCE 491 AA; 56195 MW; EE2233678B151DD1 CRC64;

Query Match 26.5%; Score 606.5; DB 16; Length 491;
Best Local Similarity 31.8%; Pred. No. 1.7e-24;
Matches 152; Conservative 100; Mismatches 187; Indels 39; Gaps 12;

Qy 3 KKLNNWIIYTTMTFTTIFLC-----LIIFPKD--TLHSELD-----DAE 46
Db 20 KRSIMLRWAPANTVFCFTFTLPTATLYQLTISSEFKEEQQLTRSDMSDVEEVLKADAP 79
Qy 47 RSSDINNLFHSKPV---KDISALDNASLGNFQEIILYDEHNKLFETSNDNTRVPEPG 103
Db 80 LNSNLTNYIATSKIQNGSEGSGLGIITGRKAFYIYD-LNHKLLYSTNRHTF----G 134
Qy 104 YEHRFYDRVIRKRYKGIEYLIKEPI--TTQDFKGYSLLIHSLNENYDNVKSLYTIALAF 161

```

```

Db 135 PQNANNEMKEIRGENPGYLVRQKIISKSTQGVVGYLOAFYDTTTHYHRIISNLLIVLLIL 194
Qy 162 GVIAITITATISYVFSQTITKPLVSLNKMTEIRRDGFQN--KLQNTNVEIDNLANTF 219
Db 195 EIVALIIVQLIGYFWANFYMKAPLKLKYGQMOEMANDPTNDFEPIEQSG-DEIEELAHVY 253
Qy 220 NEMMSQIEESFNQORQFVEDASHELRTPLQIIQGHNLNLIQWKKDPKPAVLESLSISIEE 279
Db 254 NDWMLKMKAYLEQQNRFFVSDVSHELRTPVLVDGHINLLNRWGNKDPVLDLSQASLDE 313
Qy 280 MNRILKLVLELLELTGQVNDNISSEAQTVHINDEIRSIHSLKQLHDPYQFDTLTSQVL 339
Db 314 VDRMKOMLEEMALARENVDLSSEELDCDVGKVCNRALKNFOLLHDDFEIVLD---NR 369
Qy 340 EI-----KMKPHQFEQLFLIFDINAIKYDVGNKKIKVTRLNKKOKII-EITDHGIGIPE 393
Db 370 LIYTHARISENHFEQGRILLDNNAKYSPDRKEIVITVSEDEQFVITSVSDKGIGISE 429
Qy 394 EDQDFIFDRFYRVKDSRSRSGNGGLGSLIAQKIIQLNGSGIKIKSEINKGTTTFKIIP 451
Db 430 EDINHLFERFRADKARNREIGGTGLGSLILARLAENYQGIENVSELGLGSTFTLKP 487

RESULT 7
Q07382 PRELIMINARY; PRT; 490 AA.
AC Q07382;
DT 01-JUL-1997 (TReMBLrel. 04, Created)
DT 01-JUL-1997 (TReMBLrel. 04, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Putative histidine protein kinase.
GN KINA.
OS Lactococcus lactis (subsp. cremoris) (Streptococcus cremoris).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
OX NCBI_TaxID=1359;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=MG1363;
RX MEDLINE=973116457; PubMed=9172368;
RA O'Connell-Motherway M., Fitzgerald G.F., van Sinderen D.;
RT "Cloning and sequence analysis of putative histidine protein kinases
RT isolated from Lactococcus lactis MG1363."
RL Appl. Environ. Microbiol. 63:2454-2459(1997).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=MG1363;
RX MEDLINE=20244638; PubMed=10784052;
RA O'Connell-Motherway M., van Sinderen D., Morel-Deville F.,
RA Fitzgerald G.F., Ehrlich S.D., Morel P.;
RT "Six putative two-component regulatory systems isolated from
RT Lactococcus lactis subsp. cremoris MG1363."
RL Microbiology 146:935-947(2000).
CC -!- SIMILARITY: TO OTHER PROKARYOTIC SENSORY TRANSDUCTION HISTIDINE
CC KINASES.
DR EMBL; U81166; AAC45383.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0000155; F:two-component sensor molecule activity; IEA.
DR GO; GO:0007165; P:sensory perception; IEA.
DR GO; GO:0007165; P:signal transduction; IEA.
DR InterPro; IPR003594; ATPbind_ATPase.
DR InterPro; IPR004358; Bact_sens_pr_C.
DR InterPro; IPR003660; HAMP.
DR InterPro; IPR005467; His_kinase.
DR InterPro; IPR003661; His_kinA_N.
DR Pfam; PF00672; HAMP; 1.
DR Pfam; PF02518; HATPase_C; 1.
DR PRINTS; PR00512; HisKA; 1.
DR PRINTS; PR00344; BCTRLSENSOR.
DR SMART; SM00304; HAMP; 1.

```

```

DR SMART; SM00387; HATPase c; 1.
DR SMART; SM00388; HiskA; 1.
DR PROSITE; PS50885; HAMP; 1.
DR PROSITE; PS50109; HIS_KIN; 1.
KW Kinase; Phosphorylation; Sensory transduction; Transferase.
SQ SEQUENCE 490 AA; 55884 MW; F8BDBC698D3F2039 CRC64;

Query Match      25.7%; Score 589.5; DB 2; Length 490;
Best Local Similarity 30.1%; Pred. No. 1.36-23;
Matches 143; Conservative 107; Mismatches 192; Indels 33; Gaps 11;

QY 3 KRRLNNWIIWTTMITFTVTLFLFC-----LIIFFLKD--TLHNSLDDAER----- 47
DB 20 KRSIMLRWAFANTVFCITFTFLPATLYQLTFTISSFIKEEQRLLTGSWDSAEVLEQADAP 79
QY 48 -SSDINNLFHSPKV-----KDIALDLNASLGNFOEIIYDEHNKLPFETNDNTRVVEPG 103
DB 80 LNSTNLNTYIETASKIQNGSESGSLGSIIGTRKAFYIYDLNHLKLYSTN-----RHDFG 134
QY 104 YEHRYFDRVIKRYKGYEYLIIKEPI--TQDFKGYSLLIHSLNENYDNIVKSLYIIALAF 161
DB 135 FQOQNNEMKEIQENPGYLVQRKIVSKGTQGVGYLOAFYDTTTHRIHNLVLLIL 194
QY 162 GVATITITATISYVFSTQITKPLVSLNKNMIEIRRDGFON--KIQLNTNTEEDNLANTF 219
DB 195 EIVALIVAQMIGYPMANYFMKPLEKLYGQGWASDPTNDKPIBESG-DEIEELAHVY 253
QY 220 NEMMSQIEESFNQORQFVEDASHELTPLOIIQHLNLIORWGKOPAVLEESLNTSIEE 279
DB 254 NDMMKMKAYLEQKRFVSDVSHELRTPFLAVLDGHINLLNRWKNDEPVLDESLOASLDE 313
QY 280 MNRILKVEELLELTKGVDNDISSEACTVHINDEIRSIHSLKOLHPDYQ--FDTDLTSK 337
DB 314 VDRMKRLEENLALRUENVDLSEELDCDGVKSNHALKNFQLLHEDFVIDNQIYIP 373
QY 338 NLEIKMKPHOFEQLFIDNAIKYDVKNKKIKVTKRLKNQKII-BITDHGIGIPEDQ 396
DB 374 -VHARISENHFEOGLRIILLDNNAVYSPHDRKEIIVTSEDEQFVITSVDKGIABEDI 432
QY 397 DFTDPRYRVYDKSRSGQNGGLSLTAQKLIQNGSIKIKHINKGTTFKIIF 451
DB 433 NHLFERFRADKARNRIGTGLGLPILARLAENYQGEIEVSEBELGISTFTLKF 487

RESULT 8
Q88WS3 ID Q88WS3 PRELIMINARY; PRT; 547 AA.
AC Q88WS3;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Histidine protein kinase, sensor protein (EC 2.7.3.-).
GN HPK5 OR LP_1545.
OS Lactobacillus plantarum.
OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
OC Lactobacillus.
OX NCBI_TaxID=1590;
[1]
RN RP SEQUENCE FROM N.A.
RC STRAIN=NCIMB 8826 / WCPS1;
RX MEDLINE=22480296; PubMed=12566566;
RA Kleerebezem M., Boekhorst J., van Kranenburg R., Molenaar D.,
RA Kuipers O.P., Leer R., Tarchini R., Peters S.A., Sandbrink H.M.,
RA Fiers M.W.E.J., Stiekema W., Klein Lankhorst R.M., Bron P.A.,
RA Hoffer S.M., Nierop Groot M.N., Kerkhoven R., De Vries M., Ursing B.,
RA De Vos W.M., Siezen R.J.;
RT "Complete genome sequence of Lactobacillus plantarum WCPS1.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:1990-1995(2003).
DR EMBL; AL932526; CAD63995.1; -;
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.

DR GO; GO:0000155; F:two-component sensor molecule activity; IEA.
DR GO; GO:0007165; P:signal transduction; IEA.
DR InterPro; IPR003594; ATPbind_ATPase.
DR InterPro; IPR004358; Bact_sens_pr_C.
DR InterPro; IPR003660; HAMP.
DR InterPro; IPR005467; His_kinase.
DR InterPro; IPR003661; His_kin_N.
DR Pfam; PF00672; HAMP; 1.
DR Pfam; PF02518; HATPase_c; 1.
DR Pfam; PF00512; HiskA; 1.
DR PRINTS; PR00344; BCTRLSENSOR.
DR PROSITE; PS50885; HAMP; 1.
DR PROSITE; PS50109; HIS_KIN; 1.
KW Kinase; Transferase; Complete proteome.
SQ SEQUENCE 547 AA; 62061 MW; A9119EFC50634E14 CRC64;

Query Match      24.6%; Score 563.5; DB 16; Length 547;
Best Local Similarity 32.3%; Pred. No. 3.6e-22;
Matches 137; Conservative 90; Mismatches 164; Indels 33; Gaps 11;

QY 40 SELDDAERSSDINNLFHSPKVKDIALDLNASLGNFOEIIYDEHNKLPF-----ET 92
DB 89 ASVEPATTTMSERKHGVLQSKIFSD---SDLTALSRTNLAVTVYDPHGNTLYLSRKDAHEF 145
QY 93 SNDNTVRVPGYBHRVDFRVIKRYKGYEYLIIKEPI---TTQDFKGYSLLIHSLNENYDN 149
DB 146 SKNKTQV-----VVIQSGDAE-LVGRAPFIYAASNRLIGYTVTNDLAEYHS 193
QY 150 IVKSLYIIALAFGVIAITITAT--ISYVFSTQITKPLVSLNKNMIEIRRDGFONKL--QLN 206
DB 194 TTQNLWIFWITMI-TIFGATLLGYLFAAFLRLPRMKIKQTINAVNDDPQDTSRVPLDK 252
QY 207 TYVEEDNLANTNEMMSQIEESFNQORQFVEDASHELTPLOIIQHLNLIORWGKOP 266
DB 253 RN-DESLAVAVNMDLQMQRYINQQQFVEDVSHELRTPVLIQGHMELLNRWGKDDP 311
QY 267 AVLEESLNTSIEEMNRIIKLVEELLELTKGVDNDISSEACTVHINDEIRSIHSLKOLHP 326
DB 312 QVLESLSLASLSTKMQSLVQEMLDLSRAEQLEINFSIETSQVKLVQAFDFTMIHP 371
QY 327 DYQ--FDTDLTSKNLEIKMKPHOFEQLFIDNAIKYDVKNKKIKVTKRLKNQKIIIEI 384
DB 372 DFHFTDDDV-KKPVYAQIYRNHLEQILIIILLDNNAVYSSTRKEIHLSTLDYRYIEVAV 430
QY 385 TDHGIGIPEDQDFIDRVYRVYDKSRSGQNGGLSLTAQKLIQNGSIKIKHINKG 444
DB 431 QDFGEGISKDNLDRVFNRFYVDKARSDKGGNGGLSLTAQRLVGEYGHISVESVVGQ 490
QY 445 TTFK 448
DB 491 SIPT 494

RESULT 9
Q88B09 ID Q88B09 PRELIMINARY; PRT; 462 AA.
AC Q88B09;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Sensory transduction histidine kinases.
GN BAES5 OR TTF1017.
OS Thermoanaerobacter tengcongensis.
OC Bacteria; Firmicutes; Clostridia; Thermoanaerobacteriales;
OC Thermoanaerobacteriaceae; Thermoanaerobacter.
OX NCBI_TaxID=119072;
[1]
RN RP SEQUENCE FROM N.A.
RC STRAIN=MB4 / JCM 11007;
RX MEDLINE=21992816; PubMed=11997336;
RA Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,
RA Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,
RA Tan H., Chen R., Wang J., Yu J., Yang H.;

```

"A complete sequence of T. tengcongensis genome.";

```
RT Genome Res. 12:689-700(2002).
RL EMBL; AE013066; AAM24272.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0000155; F:two-component sensor molecule activity; IEA.
DR GO; GO:00007165; F:signal transduction; IEA.
DR InterPro; IPR003594; ATPbind_ATPase.
DR InterPro; IPR004358; Bact_sens_pr_C.
DR InterPro; IPR003660; HAMP.
DR InterPro; IPR005467; His_kinase.
DR InterPro; IPR003661; His_kinA_N.
DR Pfam; PF00672; HAMP; 1.
DR Pfam; PF02518; HATPase_c; 1.
DR PRINTS; PR00512; HisKA; 1.
DR PRINTS; PR00344; BCTRLSENSOR.
DR SMART; SM00304; HAMP; 1.
DR SMART; SM00387; HATPase_c; 1.
DR SMART; SM00388; HisKA; 1.
DR PROSITE; PS50885; HAMP; 1.
DR PROSITE; PS50109; HIS_KIN; 1.
KW Kinase. Complete proteome.
SQ SEQUENCE 462 AA; 52694 MW; 80EFB35BABDFCBE CRC64;

Query Match      23.9%; Score 547; DB 16; Length 462;
Best Local Similarity 31.2%; Pred. No. 2.2e-21;
Matches 148; Conservative 93; Mismatches 183; Indels 50; Gaps 13;

QY 3 KKLRNNWIVTMTFTVTFILFCLJ-----IIFPKOTLHNSLDDAERSSDIN-NLP 56
DB ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
5 KIKLPQRTIKIALLYSAVFSVLISLNASLYTLKYLYIQSMEQVSTQAEVAKKLS 64
QY 57 HSKPVKDISALDNLASLGNFOEII---IYDEHNKKL-----FETSDNTVRVE 101
DB ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
65 EAKQWMD-GALKOTAFSGIPDENIVVKIFDSKGRLLYLSKKLVAIPYNTINLEVPVKVD 123
QY 102 PGYEHRYFDRVVKRYKGIYEL--IIEKPIITQDPFGYSLIHSLENYDNIIVKSIYII-- 157
DB ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
124 E-----FD-----KDLVYLNTVFKK-----GDATFYIQVKDMKNEYAFKLLFILMF 166
QY 158 -ALAFGVTIATITATISVFSTQITKPLVLSNKNMIEIRRDGFQNKQLNTNYEIDMLA 216
DB ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
167 FADGAGIFISVFT---GVFVTKRALRPVDYMIKEVKDIDAKGLNKLKVYGEDELTRLA 223
QY 217 NTFENWMSQIEESFNQORQFVEDASHELRTPLIIOGHNLIIQRGKKDPAPVLESNLIS 276
DB ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
224 KTFNDMLDLRESFARQNRQFVSDASHELRTPISVKGIVKIDMLDRGKKDDREVLBEKIKAI 283
QY 277 IEMNRRIIKVLELLELTGQDVNDISSEAQTVDHINDEIRSRHSIKQLHPDYQFDTDLTS 336
DB ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
284 KKETLEMESLVEKLLFLAKGDDRSIKLBRESFDLKEIAEEVVREIKLIYECK--NVLK 341
QY 337 KNLEIKMKPHQEQFLFIPINAICYVKNKKIKVTKLKNQKIKIETDTHGIGIPEDQ 396
DB ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
342 ENVRINADKKLIEVRLILDNAAKYTSKNGNIEIG-SGDEAYIKVKDQIGIPEDL 400
QY 397 DFIDRFYRVKRSRSGGNGLGIISIAQKIIQLNGSGIKIKSEINKGTTFKII 450
DB ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
401 PYIFERFVRDKARSKDTGGTGLGLSTIAKIVIEBHGGVIGVKSGKGTFTVM 454
```

RESULT 10

```
Q9LAU2
ID Q9LAU2; PRELIMINARY; PRT; 500 AA.
AC Q9LAU2;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Transmembrane histidine kinase CarS.
GN CSRS.
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
```

```
OC Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MGAS166;
RX MEDLINE=99428827; PubMed=10496909;
RA Heath A., Difita V.J., Barg N.L., Engleberg N.C.;
RT "A two-component regulatory system, CarS-CarS, represses expression of
RT three Streptococcus pyogenes virulence factors, hyaluronidase,
RT capsule, streptolysin S, and pyrogenic exotoxin B.";
RL Infect. Immun. 67:5298-5305(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MGAS166;
RA Heath A.H., Betschel S.D., Low D.E., Barg N.L.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: TO OTHER PROKARYOTIC SENSORY TRANSDUCTION HISTIDINE
CC KINASES.
DR EMBL; AF095713; AAF00082.2; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0000155; F:two-component sensor molecule activity; IEA.
DR GO; GO:0007600; F:sensory perception; IEA.
DR InterPro; IPR003594; ATPbind_ATPase.
DR InterPro; IPR004358; Bact_sens_pr_C.
DR InterPro; IPR003660; HAMP.
DR InterPro; IPR005467; His_kinase.
DR InterPro; IPR003661; His_kinA_N.
DR Pfam; PF00672; HAMP; 1.
DR Pfam; PF02518; HATPase_c; 1.
DR PRINTS; PR00512; HisKA; 1.
DR SMART; SM00304; HAMP; 1.
DR SMART; SM00387; HATPase_c; 1.
DR SMART; SM00388; HisKA; 1.
DR PROSITE; PS50885; HAMP; 1.
DR PROSITE; PS50109; HIS_KIN; 1.
KW Kinase. Phosphorylation; Sensory transduction; Transferase.
SQ SEQUENCE 500 AA; 57907 MW; FLEAFD600FDBBC9F CRC64;
```

```
Query Match      23.3%; Score 533.5; DB 2; Length 500;
Best Local Similarity 32.0%; Pred. No. 1.3e-20;
Matches 141; Conservative 82; Mismatches 163; Indels 55; Gaps 14;

QY 30 IFFLKOTLHNSLDDAE-----RSSDINNLFHSPVKDISALDNLASLGNFQIYYDE 84
DB ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
80 VLYKNDKTH-LRIDDRKGSVRISERDITN-----TLDAN-----QDIYVYNI 121
QY 85 HNNKLFETSNDNTVRPEGYEHRYFDRV---IKKRYKGIYELIIEKPIITQDPFGYSL 140
DB ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
122 DKQMIFTTDNEES---SPGL-HGPGRVYHDHIEDQYRFSMTQKVYSNRTGKFGVGVQV 177
QY 141 IHSLENDNTVKSIIYIALAFGVTIATITATISVFSTQITKPLVLSNKNMIEIRRDGFQ 200
DB ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
178 FHDLGNYVYTRARLLFWLLVVELFGTSLAYLILITTRFLKPLNHLH----EVRNISE 233
QY 201 NKLQLN-----TNYEIDNLANTFNEMMSQIEESFNQORQFVEDASHELRTPLIIOGH 255
DB ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
234 NPNNLNLRSDISSGDEIEELSVIDFNMLDKLETHTKLQSRFISDVSHELRTPVAIKGHI 293
QY 256 NLIQRWGKKDPVLEESLNISIEEMNRIIKVLELLELTK-----GDVNDISSEAQT 309
DB ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
294 GLLQRWGKKDSILLESLTATAHEADRMAINIMDLDMVRVQGSFEGHQNDM-----TV- 347
QY 310 INDEIRSRHSIKQLHPDYQFDTDLTSKNLEIKMKPHQEQFLIFTDNAIKYDVKNKKI 369
DB ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
348 LKSIVTVGNFVRVLRDFFITWQSENPXTIARIYKNHFEQALMILDNVAVKSRKEKKI 407
QY 370 KVTKRLNKKQK-IIEITDGHGIGIPEEDQDFIFORFYRVDKSRSR--SQGGNGLSLIAQ 426
DB ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

```

Db 408 AINLSVTGKQEAIVRVQDKGEGISKEDIEHIFERFYRTDKSRNRTSTQAGLIGLSILKQ 467
QY 427 IIQINGSGSIKIKSEINKGTF 447
Db 468 IVDGYHLQMKVESELNEGSVF 488

RESULT 11
O87528 PRELIMINARY; PRT; 500 AA.
AC O87528;
DT 01-NOV-1998 (T-EMBLrel. 08, Created)
DT 01-NOV-1998 (T-EMBLrel. 08, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE CSRS (Putative sensory transduction histidine kinase).
GN CSRS OR SPY0337.
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314;
RN [1]_TaxID=N.A.
RC STRAIN=DLS003;
RA Levin J.C., Wessels M.R.;
RT "Identification of csrR/csrS, a genetic locus that regulates
RT hyaluronic acid capsule synthesis in group A Streptococcus.";
RL Mol. Microbiol. 30:0-0(1998).
RN [2]
RC STRAIN=SF370 / ATCC 700294 / Serotype M1;
RX MEDLINE=21192684; PubMed=11296296;
RA Faretta J.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,
RA Primeaux C., Szate C., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,
RA Qian Y., Jia H.G., Najjar F.Z., Ren Q., Zhu H., Song L., White J.,
RA Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;
RT "Complete genome sequence of an M1 strain of Streptococcus pyogenes.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).
CC -1- SIMILARITY: TO PROKARYOTE SENSORY TRANSDUCTION PROTEINS.
DR EMBL; AF082668; AAC6436.1; -.
DR EMBL; AE006498; AAK3389.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0000155; F:two-component sensor molecule activity; IEA.
DR GO; GO:0007600; P:sensory perception; IEA.
DR GO; GO:0007165; P:signal transduction; IEA.
DR InterPro; IPR003594; ATPbind ATPase.
DR InterPro; IPR004358; Bact_sens_pr_C.
DR InterPro; IPR003660; HAMP_sens_pr_C.
DR InterPro; IPR005467; His_kinase.
DR InterPro; IPR003661; His_kinA_N.
DR Pfam; PF00672; HAMP; 1.
DR Pfam; PF00512; Hiska; 1.
DR PRINTS; PR00344; BCTRLSENSOR.
DR SMART; SM00304; HAMP; 1.
DR SMART; SM00387; HATPase_c; 1.
DR SMART; SM00388; Hiska; 1.
DR PROSITE; PS00885; HAMP; 1.
DR PROSITE; PS01019; HIS_KIN; 1.
KW Kinase; Phosphorylation; Sensory transduction; Transferase;
KW Complete proteome.
SQ SEQUENCE 500 AA; 57921 MW; 1268FB7A7D2EA5B7 CRC64;

Query Match 23.2%; Score 532.5; DB 16; Length 500;
Best Local Similarity 32.0%; Pred. No. 1.4e-20;
Matches 141; Conservative 82; Mismatches 163; Indels 55; Gaps 14;

QY 30 IFELKOTLHNSLDDAE-----RSSDINNLFHSPVKVDISALDLNASLGNFQEIYYDE 84
Db 80 VLYKNDKTH-LRIDDRKGRVIRSERDITN-----TLDAN-----QDIYVYNI 121

```

```

QY 85 HNNKLPETSDNTVRVEPGYEHRYFDRV-----IKRYKGIYLLIIEKEPITTDQFKGYSLL 140
Db 122 DKQMIFTDNEES---SPGL-HGPIGRVYVHDIEQYRGFSMTQKYSNRTGKFGYVQV 177
QY 141 IHSLENDYNIKSLYIIAFAVGVIATITATISYVFSSTQITKPLVSLSNKMIERDGFQ 200
Db 178 FHDLGNYVIRARLLFWLLVVELFGTSLAYLILITTRFLKPLHNLH---EVMENISE 233
QY 201 NKLQLN-----TNYBEIDNLANTFEMMSQIESFNQOQFVEDASHELRTPLOIIOGHL 255
Db 234 NPNNLNRSDISGDBIEELSIFONMLDKLTHTKLQSRFISDVSHSLRTPVAILKGIH 293
QY 256 NLIQRMGKDPAPVLESLSNISIEEMRIIKLVEELLELTG-----GDVNDISSEAQTVH 309
Db 294 GILQRMGKDDSLLESLTATAHEADRAIMINDMLDMIRVOGSEGHQNDM-----TV- 347
QY 310 INDEIRSIHSLKQLHPDYQFDITLTSKNLEIKMKPHQEQFLIFIDNAIKYDVNKKI 369
Db 348 LEDSIETVVGNFVLRREDFIFTWQSENPKTIARIYKQHFEQALMILIDNAVYSRKEKKI 407
QY 370 KVTRLRKQKQ-IIEITDHGIGIPEDQDFIDRFYRVVDKSR--SOGNGLGLSIAQ 426
Db 408 AINLSVTGKQEAIVRVQDKGEGISKEDIEHIFERFYRTDKSRNRTSTQAGLIGLSILKQ 467
QY 427 IIQINGSGSIKIKSEINKGTF 447
Db 468 IVDGYHLQMKVESELNEGSVF 488

RESULT 12
O8P2J7 PRELIMINARY; PRT; 500 AA.
AC O8P2J7;
DT 01-OCT-2002 (T-EMBLrel. 22, Created)
DT 01-OCT-2002 (T-EMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE CovS.
GN COVS OR SPYM18_0329.
OS Streptococcus pyogenes (serotype M18).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=186103;
RN [1]_TaxID=N.A.
RC STRAIN=NGAS8232 / Serotype M18;
RX MEDLINE=21927593; PubMed=11917108;
RA Smoot J.C., Barbican K.D., Van Gompel J.J., Smoot L.M., Chaussee M.S.,
RA Sylva G.L., Sturdevant D.E., Ricklefs S.M., Porcella S.F.,
RA Parkins L.D., Beres S.B., Campbell D.S., Smith T.M., Zhang Q.,
RA Kapur V., Daly J.A., Veasy L.G., Musser J.M.;
RT "Genome sequence and comparative microarray analysis of serotype M18
RT group A Streptococcus strains associated with acute rheumatic fever
RT outbreaks.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:4668-4673(2002).
DR EMBL; AB009978; AAL97084.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0000155; F:two-component sensor molecule activity; IEA.
DR GO; GO:0007165; P:signal transduction; IEA.
DR InterPro; IPR003594; ATPbind ATPase.
DR InterPro; IPR004358; Bact_sens_pr_C.
DR InterPro; IPR003660; HAMP.
DR InterPro; IPR005467; His_kinase.
DR InterPro; IPR003661; His_kinA_N.
DR Pfam; PF00672; HAMP; 1.
DR Pfam; PF02518; HATPase_c; 1.
DR Pfam; PF00512; Hiska; 1.
DR PRINTS; PR00344; BCTRLSENSOR.
DR PROSITE; PS00885; HAMP; 1.
DR PROSITE; PS01019; HIS_KIN; 1.
KW Complete proteome.
SQ SEQUENCE 500 AA; 57907 MW; 1268FCOA7D2EA5B7 CRC64;

```



```
DR GO; GO:000155; F:two-component sensor molecule activity; IEA.
DR GO; GO:0007165; P:signal transduction; IEA.
DR InterPro; IPR003594; ATPbind_ATPase.
DR InterPro; IPR004358; Bact_sens_pr_C.
DR InterPro; IPR003660; HAMP.
DR InterPro; IPR005467; His_kinase.
DR InterPro; IPR003661; His_kin_N.
DR Pfam; PF00672; HAMP; 1.
DR Pfam; PF02518; HATPase_c; 1.
DR Pfam; PF00512; HSKA; 1.
DR PRINTS; PR00344; BCTRLSENSOR.
DR SMART; SM00304; HAMP; 1.
DR SMART; SM00387; HATPase_c; 1.
DR SMART; SM00388; HSKA; 1.
DR PROSITE; PS00885; HAMP; 1.
DR PROSITE; PS50109; HIS_KIN; 1.
DR KINASE; Complete proteome.
SQ SEQUENCE 500 AA; 57963 MW; 1268PB6E383AE0B7 CRC64;

Query Match 22.9%; Score 523.5; DB 16; Length 500;
Best Local Similarity 31.7%; Pred. No. 4.3e-20;
Matches 140; Conservative 82; Mismatches 164; Indels 55; Gaps 14;

QY 30 IFFLKLDTLHNSLDDAE-----RSSSDINNLFSKPKVKDISALDLNASLGNFOBIIIYDE 84
DB 80 VLYKNDKTH-LRIDDRKGSVRISERDITN-----TLDAN-----QDIYVYNI 121
QY 85 HNNKLFSTNDNTVVRVPGVEHYRDRV----IKKRYKGYEYLIIEKEPIITQDFKGYSL 140
DB 122 DKQIMFTTNEES---SPGL-HGPIGRVYHDHIEDYRGFSMTQKVYSNRTGKFGVGVQV 177
QY 141 IHSLENDYVKSIIYIITAFGLVATITATISVVFSTQITKPLVSLSNKMKIEIRRDGFQ 200
DB 178 FHDLGNYIVRARLFWLLVVELFGTSLAYLIILITTRFLKPLNLH-----EVMRNISE 233
QY 201 NKQLQN-----TNYEEDINLANTFNEMWSQIESFNQORQVEDASHELRTPLQIQHGL 255
DB 234 NPNNLNLRSIDSSGDEIEELSVPDNMLDKLEHTKLSRFISDVSHSLTPTVAIKGHI 293
QY 256 NLQRWKQKDPVLEESLNTSIEEMNRIKLVEELLETK-----GDVNDISSEAOQTVH 309
DB 294 GLLQWKGKPSDILLESITATATAHEADRAIMINDMLDIRVQGSFEGHQNDM-----TV- 347
QY 310 INDEIRSRHSLKOLHPDYQFDITLTSKNLEIKMKPHQFEQLFLIFIDNAIKYDVKNKKI 369
DB 348 LEDSIETVGVNFRVLRDEFIFTQSENPKTARIYKNHFEQALMILIDNAVKYSRKEKKI 407
QY 370 KVTRLKNKQK-IIEITDHGIGIPEEQDFTFRFYRVDSKRSR--SQGGNGLGLSLTAOK 426
DB 408 AINLSVTGKQEAIVRVQDKGEGISKEDIEHIFERFYTDKSRNRTSTQAVLGIGLSILQ 467
QY 427 IIQNGSGSIKIKSINKGTTTF 447
DB 468 IVDGYHLQMKVSELSNEGVSF 488

RESULT 15
Q8E3T5 PRELIMINARY; PRT; 501 AA.
AC Q8E3T5;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN GBS1671.
OS Streptococcus agalactiae (serotype III).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=216495;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NEW316 / Serotype III;
RX MEDLINE=22242508; PubMed=12354221;

RA Glaser P., Rusniok C., Buchrieser C., Chevalier F., Frangeul L.,
RA Meadek T., Zouine M., Couve E., Lalioui L., Poyart C., Trieu-Cuot P.,
RA Kunst F.;
RT "Genome sequence of Streptococcus agalactiae, a pathogen causing
RT invasive neonatal disease.";
RL Mol. Microbiol. 45:1499-1513(2002).
DR EMBL; AL766852; CAD47330.1; -.
DR Sagaliet; gbs1671; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:000155; F:two-component sensor molecule activity; IEA.
DR GO; GO:0007165; P:signal transduction; IEA.
DR InterPro; IPR003594; ATPbind_ATPase.
DR InterPro; IPR004358; Bact_sens_pr_C.
DR InterPro; IPR003660; HAMP.
DR InterPro; IPR005467; His_kinase.
DR InterPro; IPR003661; His_kin_N.
DR Pfam; PF00672; HAMP; 1.
DR Pfam; PF02518; HATPase_c; 1.
DR Pfam; PF00512; HSKA; 1.
DR PRINTS; PR00344; BCTRLSENSOR.
DR SMART; SM00304; HAMP; 1.
DR SMART; SM00387; HATPase_c; 1.
DR SMART; SM00388; HSKA; 1.
DR PROSITE; PS00885; HAMP; 1.
DR PROSITE; PS50109; HIS_KIN; 1.
DR KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 501 AA; 57482 MW; 42C91E46581F40DE CRC64;

Query Match 22.5%; Score 516; DB 16; Length 501;
Best Local Similarity 27.7%; Pred. No. 1.1e-19;
Matches 134; Conservative 106; Mismatches 169; Indels 74; Gaps 12;

QY 20 VTIFLFLIIIFPL-----KDTLHNSL-----DAERSSSDINNLPH 57
DB 23 VFLFSLFTVFSVLVYTSATRVLRKINVGSRLEKTRVRLSQANSSLTSDILEILY 82
QY 58 SKPVKD-----ISALDLNASLGNFOBIIIYDEHNNKLFSTNDNTVVRVPGVE 105
DB 83 NQVFADDIYPHKRQNGIVRTGESIDSILYANQEMTLYDVNRKPVF-----STLRTQMPTI 137
QY 106 HRYFDRVIVKRYKGYEYLIIEKEPIITQ---DFKGYSLLSLENDYVKSIIYIITAFGL 162
DB 138 GSKMGKVIISKVADMEGFGVTGKAIYSQTKQLGLGVQVIFNGLGRYYSMKQNIIVFLIMME 197
QY 163 VIATITATISVVFSTQITKPLVSLSNKMKIEIRRDGFONKQLNT---NYEEDINLANTF 219
DB 198 VLGTVALVINSATKRIVRPVKNLHDLMHQISEN--PSNLEIRSKVRSEDEIGELSRIF 255
QY 220 NEMWSQIESFNQORQVEDASHELRTPLQIQHGLNLQRWKQKDPVLEESLNTSIEE 279
DB 256 DGMLDQLEDYTRRSQFISDVSHSLTPTVAVVKGHIGLLQWKGKDPPEILEESLAAAYHE 315
QY 280 MARIILKVEEL-----LELTGKVNDISSEAOQTVHINDEIRSRHSLKOLHPDYQF 330
DB 316 ADRLSLMINDMLNMRVQGSLELHQDEVTLSSISV-----ENFRILRDRDQF 366
QY 331 DTDLTSLNLEI-----KMKPHQFEQLFLIFIDNAIKYDVKNKKIKVTRLKNKQKIIETD 386
DB 367 ----IFENNISDIVMGKIYKHFEQALMILIDNAIKYSPSYKSVSVLSVDNDPATVVKD 423
QY 387 HGIGIPEEQDFTFRFYRVDSKRSR--SQGGNGLGLSLTAQKIIQNGSGSIKIKSINKG 444
DB 424 KGEISDEIDIEFTFRFYTDKSRNRESTQAGLIGLGVFKQIMDAYHLKVIDIKSELNOG 483
QY 445 TTF 447
DB 484 TEF 486

Search completed: October 4, 2004, 18:09:27
Job time : 106.624 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 4, 2004, 17:51:42 ; Search time 20.194 Seconds
(without alignments)
1162.900 Million cell updates/sec

Title: US-09-006-627-2

Perfect score: 2291

Sequence: 1 MTKRLRNWIIVTMTTFV.....GGSIKIKSEINKGTTFKIIF 451

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	471	20.6	454	1 YKOH_BACSU	O34638 bacillus su
2	405.5	17.7	611	1 YCG_BACSU	Q45614 bacillus su
3	385.5	16.8	589	1 RESE_BACSU	P35164 bacillus su
4	383	16.7	473	1 YCLK_BACSU	P94414 bacillus su
5	368.5	16.1	579	1 PHOR_BACSU	P23545 bacillus su
6	360.5	15.7	480	1 CUSS_ECOLI6	O8fk37 escherichia
7	358	15.6	444	1 CIAH_STRPN	Q54955 streptococc
8	353.5	15.4	482	1 CUSS_ECO57	Q8xb74 escherichia
9	349.5	15.3	480	1 CUSS_ECOLI	P77485 escherichia
10	340	14.8	497	1 SILS_SALTY	Q9zhd4 salmonella
11	335	14.6	565	1 MTRB_MYCPA	O93cb7 mycobacteri
12	331.5	14.5	918	1 BARA_SHIFL	P59342 shigella fl
13	330.5	14.4	918	1 BARA_ECOLI	P28607 escherichia
14	328	14.3	567	1 MTRB_MYCHO	P59963 mycobacteri
15	328	14.3	567	1 MTRB_MYCTU	Q50496 mycobacteri
16	319	13.9	535	1 AFQ2_STRCO	O49493 streptomyce
17	317.5	13.9	562	1 MTRB_MYCLE	O9cc11 mycobacteri
18	300.5	13.1	475	1 C2CS_ALCEU	O44007 alcaligenes
19	297	13.0	467	1 BRES_ECOLI	P30847 escherichia
20	289	12.6	414	1 CUTS_STRCO	Q03757 streptomyce
21	283.5	12.4	596	1 CHVG_AGR75	O07737 agrobacteri
22	275	12.0	663	1 DSPA_SYNY3	P20169 synecocyst
23	270.5	11.8	431	1 PHOR_ECOLI	P08400 escherichia
24	265.5	11.6	457	1 CPXA_ECOLI	P08336 escherichia
25	265	11.6	450	1 SEX3_MYCTU	Q1155 mycobacteri
26	263.5	11.5	451	1 CSSS_BACSU	O32193 bacillus su
27	263	11.5	464	1 IRLS_BURPS	O31396 burkholderi
28	262.5	11.5	577	1 CHVG_RHIME	P72292 rhizobium m
29	260.5	11.3	431	1 PHOR_KLEPN	P45608 klebsiella
30	258	11.3	452	1 YEDV_ECOLI	P76339 escherichia
31	257.5	11.2	431	1 PHOR_SHIDY	P45609 shigella dy
32	254.5	11.1	466	1 PCOS_ECOLI	Q47457 escherichia
33	251.5	11.0	447	1 VANS_ENTPA	Q47745 enterococcu

ALIGNMENTS

RESULT 1

ID	YKOH_BACSU	STANDARD;	PRT;	454 AA.
AC	O34638;			
DT	28-FEB-2003 (Rel. 41, Created)			
DT	28-FEB-2003 (Rel. 41, Last sequence update)			
DT	10-OCT-2003 (Rel. 42, Last annotation update)			
DE	Hypothetical sensor-like histidine kinase ykoH (EC 2.7.3.-).			
GN	YKOH OR BSU13260.			
OS	Bacillus subtilis.			
OC	Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.			
OX	NCBI_TaxID=1423;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=168;			
RT	Devine K.M.;			
RT	"Sequence of the Bacillus subtilis genome between xlyA and ykoR.";			
RL	Submitted (NOV-1997) to the EMBL/GenBank/DBSJ databases.			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=168;			
RT	Medline=98044033; PubMed=9384377;			
RA	Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G., Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S., Borries R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S., Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M., Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A., Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T., Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D., Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N., Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G., Guiseppe G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A., Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L., Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C., Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M., Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V., Lee S.M., Levine A., Liu H., Masuda S., Maubel C., Medigue C., Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M., Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H., Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M., Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S., Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y., Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F., Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K., Takeuchi M., Tanakoshi A., Tanaka T., Terpetra P., Tognoni A., Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A., Viari A., Wambuit R., Wedler E., Wedler H., Weitzenecker T., Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K., Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.; "The complete genome sequence of the Gram-positive bacterium Bacillus subtilis"; Nature 390:249-256 (1997).			

CC - FUNCTION: Could be member of the two-component regulatory system ykoH/ykoG. Potentially phosphorylates ykoG.

CC - SUBCELLULAR LOCATION: Integral membrane protein (Potential).

or send an email to license@sib.ch).

```
CC EMBL; AE016757; AAN79131.1;
DR InterPro; IPR003594; ATPbind_ATPase.
DR InterPro; IPR004358; Bact_sens_dr_C.
DR InterPro; IPR003860; HAMF.
DR InterPro; IPR003661; His_kinA_N.
DR InterPro; IPR005467; His_kinase.
DR InterPro; IPR006290; Metal_his_kin.
DR Pfam; PF00672; HAMF; 1.
DR Pfam; PF02518; HATPase_c; 1.
DR Pfam; PF00512; HisKA; 1.
DR PRINTS; PR00344; BCTRLSNSOR.
DR SMART; SMO0304; HAMF; 1.
DR SMART; SMO0387; HATPase_c; 1.
DR SMART; SMO0388; HisKA; 1.
DR TIGRFAMs; TIGR01386; cztS_sils_cps; 1.
DR PROSITE; PS50885; HAMF; 1.
DR PROSITE; PS50109; HIS_KIN; 1.
KW Sensory transduction; Transferase; Kinase; Transmembrane;
KW Inner membrane; Phosphorylation; Copper; Complete proteome.
FT DOMAIN 1 15 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 16 36 POTENTIAL.
FT DOMAIN 37 186 PERIPLASMIC (POTENTIAL).
FT TRANSMEM 187 207 POTENTIAL.
FT DOMAIN 208 480 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 207 260 HAMF.
FT DOMAIN 268 480 HISTIDINE KINASE.
FT MOD_RES 271 271 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
SQ SEQUENCE 480 AA; 53776 MW; 34C2D8A8A38F16E CRC64;

Query Match 15.7%; Score 360.5; DB 1; Length 480;
Best Local Similarity 23.4%; Pred. No. 8.1e-13;
Matches 114; Conservative 102; Mismatches 223; Indels 49; Gaps 11;

QY 1 MTKRKLNNWIIVTMTITFTVI-----FLCLIIIFLKTLLNSELDDAERSSDINN 54
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1 MVSKPFPQFSLATRLTFISLIATTAFAFAFWMIHVKVHFADQINDLKEISATLER 60
QY 55 -LEFSKPKVKDISALDNLASLGNFQIIIVDEHNKLFETSDNTVTRVEPGY-EHRYF--D 110
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 61 VLNHPDETQARRLMTLEDIVSGYSNVL-----SLADSHGKTVYHSPGAPDREFTRD 113
QY 111 RVIKRYKIGIYLIKEPITTPQPKGYSLIHSLENYDNIV-----KSLYIITALAPGV 163
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 114 AIPDKAOGGEVYLLSGPTWMPGCHGHMEHSNWRMLNLPVGLVDGKPIYLYIALSI 173
QY 164 -----IATITATISYVFSTQITK---PLVSLNKNMIEIRRDGFQNKQL 204
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 174 DFHLHYINDLMNKLMTASVISILVFTVLLAVHKHAPIRSVSRQIQNITSKDLVDRLD 233
QY 205 LNTNVEEDNLANTFNEMMSIESFNQORQFVEDASHELTPLO--IIQHLNLIQRWG 262
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 234 POTVPIELEQLVLFNHNMIERIEVFTQSFNSADIAHEIRTPITNLITOTEIALSQSRS 293
QY 263 KKDPAVLEESNISIEENNRRIKLVLEELELTKGVDNISSEAOVTVNDIERSKHSK 322
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 294 QKE---LEDVLYSNELELTRAKVSDMLFLAQNQNLIEKKWMLNLADEVGVDFPFE 350
QY 323 QLHPDYQPDTLTSKNLEIKMKPHQFQELFIIDNAIKYDVQNKIKVKTRLKXKQKII 382
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 351 ALAEDRGVELRVFGDECQVAGDPLMLRLNLLSNALRYPTGETIIVRCQTVDLHVQV 410
QY 383 EITDHGIGIPEDQDFIDRFRYVDKSRSGQGNGLSLAQKIIQLNGSIIKSEIN 442
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 411 TVENPGTPIAEHLRLPDRFRYVDPSRQKRGEGSGIGIALIVKSIIVAHKGTVAVTSDV- 469
QY 443 KGTTFKII 450
DQ : : : : :
DB 470 RGTDFVII 477
```

RESULT 7

```
CIAH STRPN
ID CIAH STRPN STANDARD; PRT; 444 AA.
AC Q54955;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Sensor protein ciah (EC 2.7.3.-).
GN CIAH OR SP0799 OR SP0708.
OS Streptococcus pneumoniae, and
OC Streptococcus pneumoniae (strain ATCC BAA-255 / R6).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313, 171101;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC BAA-255 / R6;
RX MEDLINE=94344038; PubMed=8065267;
RA Guenzi E., Gasc A.M., Sicard M.A., Hakenbeck R.;
RT "A two-component signal-transducing system is involved in competence
and penicillin susceptibility in laboratory mutants of Streptococcus
pneumoniae.";
RL Mol. Microbiol. 12:505-515(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC BAA-334 / TIGR4;
RX MEDLINE=21357209; PubMed=11463916;
RA Tettelin H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,
Peterson S., Heidelberg J., DeBoy R.T., Haft D.H., Dodson R.J.,
Durkin A.S., Gwinn M., Kolonay J.F., Nelson W.C., Peterson J.D.,
Umayan L.A., White O., Salzberg S.L., Lewis M.R., Radune D.,
Holtzapple E., Khouri H., Wolf A.M., Uutterback T.R., Hansen C.L.,
McDonald L.A., Feldblyum T.V., Angiuoli S., Dickinson T., Hickey E.K.,
Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.,
Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;
RT "Complete genome sequence of a virulent isolate of Streptococcus
pneumoniae.";
RL Science 293:498-506(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC BAA-255 / R6;
RX MEDLINE=21429243; PubMed=11544234;
RA Hoskins J., Albom W.E. Jr., Arnold J., Blaszcak L.C., Burgett S.,
DeHoff B.S., Estrem S.T., Fritz L., Fu D.-J., Fuller W., Geringer C.,
Gilmour R., Glass J.S., Khoja H., Kraft A.R., Lagace R.E.,
Leblanc D.J., Lee L.N., Lefkowitz E.J., Lu J., Matsushima P.,
McAhren S.M., McHenry M., McLeaster K., Mundy C.W., Nicas T.I.,
Norris F.H., O'Garra M., Peery R.B., Robertson G.T., Rockey P.,
Sun P.-M., Winkler M.E., Yang Y., Young-Bellido M., Zhao G.,
Zook C.A., Baltz R.H., Jaskunas S.R., Rostek P.R. Jr., Skatrud P.L.,
Glass J.I.;
RT "Genome of the bacterium Streptococcus pneumoniae strain R6.";
RL J. Bacteriol. 183:5709-5717(2001).
CC -!- FUNCTION: Member of the two-component regulatory system ciah/ciar.
Involved in early steps of competence regulation and in penicillin
susceptibility. Probably phosphorylates ciar.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: Contains 1 histidine kinase domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.ebi.ac.uk/ebis-sib.ch).
CC or send an email to license@sib.ch.
CC
CC EMBL; X77249; CAA54466.1; -.
DR EMBL; AE007386; AAK74936.1; -.
DR EMBL; AE008447; AAK99812.1; -.
DR PIR; D97960; D97960.
DR PIR; G95092; G95092.
DR PIR; S49545; S49545.
DR TIGR; SP0799; -.
```

[illegible]

```
QY 1 MTKRLNNWIIWTVITVTI-----FLCIIIIFFLKOTLHNSLDDAERSSSDINN 54
Db 1 MYSKFPORPFSLATRLTTFISLATIAAFAPAFAMIHVSVKVHFARQDINDLKEISATLER 60
QY 55 -LPHSKPVKDISALDINASLGNFQEIILYDEHNKLFETSNDNTVYRVEGY-EHRYF--D 110
Db 61 VLNHPEDEQARLMTELDIVSGYNNLI-----SLADSHGKTVYHSGAPDIREFARD 113
QY 111 RVIKRYKGIIEYLIKEPITTDQFKGYSLIHSLENYDNIV-----KSLYIIIALAF 161
Db 114 AIPDKDARGEVLLSGPTMMMPGHGHMEHS--NWRMISLPVGLVDGKPIYLYIAL 171
QY 162 GV-----IATITATISVFTQTK---PLVLSNKKMEIRRDGFONK 202
Db 172 SIDFHLVINDLNNKLIMTASVISLIVFVLAVHKGHAPIRSVRSQITNSKDLVR 231
QY 203 LQNTNYYEIDNLANTFEMMSQIBESFNQORQFVEDASHELRTPLO--LIQHLNLLOR 260
Db 232 LDPQVPIEQLVLSFNHMERIEDVTRQNFADIAHEIRTPITNLITQTEALSQS 291
QY 261 WKKDPAVLESLSNISIEEMRIIKLVBELLELTGKGVNDISSEATVHINDEIRSIHS 320
Db 292 RSQKE--LELVLYSNEELTRMAKWSMDLFLAQADNNQLIPEKKMLNLADVEGKVDF 348
QY 321 LKQLHPDQFDTLTSKLEIKMKHPQEQFLIFIDNAIKVDVKNKKIKVTRUNKQK 380
Db 349 FEALAEGRVQLQFVQDSCVAGDPLMLRRALSNNLSNALRYTPPGEAIVVRCQTVDLV 408
QY 381 IIEITDHGIGPEEDQDFRFRVYDKSRSSQGGNGLSIAQKIQLNGSGSIKISE 440
Db 409 QVIVENPTPIAEPHLRFLRFDYRQKRGEGSGIGLAIVKSVVAHKGTVAVTSN 468
QY 441 INKGTTFKII 450
Db 469 A-RGTRFVIV 477

RESULT 9
CUSS_ECOLI STANDARD; PRT; 480 AA.
AC P77A85; Q9R7T9; Q9R7U0;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Sensor kinase cusB (EC 2.7.3.-).
GN CUSS OR B0570.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN (1)
RP SEQUENCE FROM N.A., AND FUNCTION.
RC STRAIN=K12 / DH5-alpha;
RX MEDLINE=20461235; PubMed=11004187;
RA Munson G.P., Lam D.L., Outten F.W., O'Halloran T.V.;
RT "Identification of a copper-responsive two-component system on the
RT chromosome of Escherichia coli K-12.";
RL J. Bacteriol. 182:5864-5871 (2000).
RN (2)
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474 (1997).
RN (3)
RP SEQUENCE FROM N.A.
RA Chung E., Allen E., Araujo R., Aparicio A., Davis K., Duncan M.,
RA Federpriel N., Hyman R., Kaiman S., Komp C., Kurdi O., Lew H., Lin D.,
RA Namath A., Oefner P., Roberts D., Schramm S., Davis R.W.;
```

```
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=97061202; PubMed=8905232;
RA Oshima T., Alba H., Baba T., Fujita K., Hayashi K., Honjo A.,
RA Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,
RA Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,
RA Mori H., Motomura K., Nakamura Y., Nashimoto H., Nishio Y., Saito N.,
RA Saueki G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,
RA Yano M., Horiuchi T.;
RT "A 718-Kb DNA sequence of the Escherichia coli K-12 genome
RT corresponding to the 12.7-28.0 min region on the linkage map.";
RL DNA Res. 3:137-155 (1996).
RN [5]
RP FUNCTION IN COPPER HOMEOSTASIS.
RC STRAIN=K12;
RX MEDLINE=21319120; PubMed=11399769;
RA Outten F.W., Huffman D.L., Hale J.A., O'Halloran T.V.;
RT "The independent cue and cus systems confer copper tolerance during
RT aerobic and anaerobic growth in Escherichia coli.";
RL J. Biol. Chem. 276:30670-30677 (2001).
RN [6]
RP PROBABLE FUNCTION IN SILVER HOMEOSTASIS.
RC STRAIN=K38;
RX MEDLINE=21178899; PubMed=11283292;
RA Franke S., Grass G., Nies D.H.;
RT "The product of the ybdE gene of the Escherichia coli chromosome is
RT involved in detoxification of silver ions.";
RL Microbiology 147:985-972 (2001).
CC -!- FUNCTION: Member of the two-component regulatory system cusC/cusR.
CC Copper ion sensor. Could also be a silver ion sensor. Probably
CC activates cusR by phosphorylation.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
CC (Probable).
CC -!- MISCELLANEOUS: The cus system plays an important role in copper
CC tolerance under anaerobic growth and, under extreme copper stress,
CC in aerobic growth.
CC -!- SIMILARITY: Contains 1 HAMP domain.
CC -!- SIMILARITY: Contains 1 histidine kinase domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF245661; AAF70176.1; -
CC EMBL; AE000162; AAC73671.1; -
CC EMBL; U82598; AAB40768.1; -
CC EMBL; D90699; BAA35204.1; ALT_INIT.
CC EMBL; D90700; BAA35210.1; ALT_INIT.
CC PIR; H64789; H64789.
CC HSSP; P02933; 1BXD.
CC EcoGene; EGI3642; CUS.
CC InterPro; IPR003594; ATPbind_ATPase.
CC InterPro; IPR004358; Bact_sens_pr_C.
CC InterPro; IPR003660; HAMP.
CC InterPro; IPR003661; His_kinA_N.
CC InterPro; IPR005467; His_kinase.
CC InterPro; IPR006290; Metal_his_kin.
CC Pfam; PF00672; HAMP; 1.
CC Pfam; PF02518; HATPase_c; 1.
CC Pfam; PF00512; Hiska; 1.
CC PRINTS; PR00344; BCTRSENSOR.
CC SMART; SM00304; HAMP; 1.
CC SMART; SM00387; HATPase_c; 1.
CC TIGRFAMs; TIGR01386; cztS_silS_copS; 1.
CC PROSITE; PS00885; HAMP; 1.
CC PROSITE; PS0109; HIS_KIN; 1.
```



```

KW Sensory transduction; Transferase; Kinase; Transmembrane;
KW Inner membrane; Phosphorylation; Copper; Complete proteome.
FT DOMAIN 1 15 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 16 36 POTENTIAL.
FT DOMAIN 37 186 PERIPLASMIC (POTENTIAL).
FT TRANSMEM 187 207 POTENTIAL.
FT DOMAIN 208 480 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 208 260 HAMP.
FT DOMAIN 268 480 HISTIDINE KINASE.
FT MOD_RES 271 271 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
SQ SEQUENCE 480 AA; 53737 MW; B3D78C0810241519 CRC64;

Query Match
Best Local Similarity 23.28; Pred. No. 3.2e-12; Length 480;
Matches 113; Conservative 100; Mismatches 225; Indels 49; Gaps 11;

QY 1 MTRKLNNNIIVTMTITFTI-----FLFCLIIIFFLKDTLHNSLDDAERSSSDINN 54
DB 1 MVSQPFQPSLATRLTFFISLATIAFAFAFAMIHVSQVHFAEQDINDLKEISATLER 60
QY 55 -LFHSKPKVDISALDNLASLGNQFIIIDYHNNKLPETSDNTVTRVEPGY-EHRYF--D 110
DB 61 VLNHPDETQARRLMTLEDIVSGYSNVLI-----SLADSQKGTVVYHSPGADIREFTD 113
QY 111 RVIKRYKGIEYLIIEPIITTDQPKGYSLIHSLNVDIV-----KSLYIIIALAFGV 163
DB 114 AIPDKAQGGEVLLSGTMMPGHGHGHEHNSWRMINLPVGLVDGKPIYLYIALSI 173
QY 164 -----IATITATISYVFSTQITK-----PLVSLSNKMKIEIRRDGFQNKLO 204
DB 174 DFHLHYINDLNNKLIMTASVISILVFI VLLAVHKGHAPIRSVSRQINITSKDLVDRLD 233
QY 205 LNTYIEIDNLANTFNEMSQIEBSFNQOQFVEDASHELRTPLQ--IIQCHLNLTQW 262
DB 234 POTVPIEQLVLSFNMIERIEDVFTQSFNSADIAHEIRTPITNLITQTEIALSQSRS 293
QY 263 KQDPAVLESLNISIREMNIIRKIVELLETGKGVNDVDSSEAOQTVHNDIEIRSHSLK 322
DB 294 QKE---LEDVLYSNLELTMAKVSNDLFLAQNQLIPEKKMLNLADEVKVPDFPE 350
QY 323 QLHPDYQFDTLTSKNLEIRKPKHQFQLFLIFIDNAIKYDVKNKKIKVTRLNKKQKII 382
DB 351 ALAEDRGVELRFVQKQVAGDPLMLRRLSNLLSNALRYTPGTETIVVRCQTVHLVQV 410
QY 383 EITDHGIGIPEEDQDFIDRFYRVKRSRQSGNGGLGSLIAOKIQLNGSGSIKISEIN 442
DB 411 IVENPGTFIAPEHLPRLDFRFDYRDPQRQKGGSGIGLAIVKSVIVVAHKGTVAVTSDA- 469
QY 443 KGTTFKI 449
DB 470 RGTFRVI 476

RESULT 10
SILS_SALT
ID_SILS_SALT STANDARD; PRT; 497 AA.
AC Q9ZHD4;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Probable sensor kinase sils (EC 2.7.3.-).
GN SILS.
OS Salmonella typhimurium.
OG Plasmid pMG101.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=602;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99128056; PubMed=9930866;
RA Gupta A., Matsui K., Lo J.-F., Silver S.;
RT "Molecular basis for resistance to silver cations in Salmonella.";
RL Nat. Med. 5:183-188(1999).

```

```

CC -!- FUNCTION: COMPONENT OF THE SIL CATION-EFFLUX SYSTEM THAT CONFERS
CC RESISTANCE TO SILVER. PROBABLE MEMBER OF A TWO-COMPONENT
CC REGULATORY SYSTEM SILS/SILR. MAY ACTIVATE SILR BY PHOSPHORYLATION.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
CC (Probable).
CC -!- SIMILARITY: Contains 1 HAMP domain.
CC -!- SIMILARITY: Contains 1 histidine kinase domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF067954; AAD11744.1; -.
CC InterPro; IPR003594; ATPbind_ATPase.
CC InterPro; IPR004358; Bact_sens_pr_C.
CC InterPro; IPR003660; HAMP.
CC InterPro; IPR003661; His_kinase.
CC InterPro; IPR005467; His_kinase.
CC InterPro; IPR006290; Metal_his_kin.
CC Pfam; PF00672; HAMP; 1.
CC Pfam; PF02518; HATPase_c; 1.
CC Pfam; PF00512; HisKA; 1.
CC PRINTS; PR00344; BCTRLSENSOR.
CC SMART; SM00304; HAMP; 1.
CC SMART; SM00387; HATPase_c; 1.
CC SMART; SM00388; HisKA; 1.
CC TIGRFAMs; TIGR01386; czts_sils_cops; 1.
CC PROSITE; PS50885; HAMP; 1.
CC PROSITE; PS50109; HIS_KIN; 1.
CC Sensory transduction; Transferase; Kinase; Transmembrane;
KW Inner membrane; Phosphorylation; Plasmid.
FT DOMAIN 1 15 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 16 36 POTENTIAL.
FT DOMAIN 37 186 PERIPLASMIC (POTENTIAL).
FT TRANSMEM 187 207 POTENTIAL.
FT DOMAIN 208 497 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 208 261 HAMP.
FT DOMAIN 269 487 HISTIDINE KINASE.
FT MOD_RES 272 272 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
SQ SEQUENCE 497 AA; 56246 MW; C5F19D6DCL1A0D96 CRC64;

Query Match
Best Local Similarity 23.8%; Pred. No. 1.1e-11; Length 497;
Matches 114; Conservative 92; Mismatches 204; Indels 70; Gaps 13;

QY 13 VITMTITFTVIFLECLIIIFFLKDT-LHNSLDD--DAERSSSDINNLFHSPKPKVDISALD- 68
DB 16 LTFIFSLSTILAFIAPTWFMLHSEVKEHFAEQDVSDDQOISTLSRILQSPADPEKKVSK 75
QY 69 LNASLGNFOEI--IYDEHNNKLFETSDNTVTRVEPGYHRYFDRVIKRYGIEYLIK 126
DB 76 IKESIASYRNVALLLNPRGEVLYSSAQGAALR--PAVNSADPSE--HSRARDVFLTWVE 131
QY 127 EPIITQD-----FKGYSLIHSLNRYDNIVKSLYI 156
DB 132 DTARAMDTGSGMKMETYRIIASSGQATFOGKQQNVYMLTGLSINPH-LHYLDALKKNL-- 188
QY 157 IALAFGVIAITITATISYVFTSQITKPLVSLSNKMKIEIRRDGFQNKQLQNTVEEDNLA 216
DB 189 --IAIAVSLIIVLIIRIAVRQGHLPFRNVSNKINITSENLDARLETPTRVPIEQLV 246
QY 217 NTFNEMKSOIERSFNQOQFVEDASHELRTPLQ--IIQCHLNLIQWKKDPAVLBSLN 274
DB 247 ISFNEMIGKIEDVFTQANFSADIAHEIRTPITNLVTOETIALSQDRTQKE---LEDVLY 303
QY 275 ISIEENNRRIKVLVELELTGKGVNDVDSSEAOQTVHNDIEIRSHSLKQLHPDYQFDTL 334
DB 304 SSLEENNRMTKVMVDMFLAQADNNQL-----IPDRVRFDLQSQNSLKVRVRFSEAL 355

```



```
QY 335 TSKNLEIKM-----KPHQEQFLFIPIDNAIKYDVNKKIKVTKLKNKQKII 382
Db 356 GKPTPIILLKFGMPCLVEGDPQFRRAINNLLSNALRYTPEGQAITVSIREQESFFDL 415
QY 383 EITDHGIGIPEDQDPIFDRFVYDKSRSGGNGLSIAQKIQLNGSGSIKIKSEIN 442
Db 416 VIENFGKPIPEHLSRLFDRFVRDPSRQKGGGIGLAIVKSIVEAHGHRVQVESDVH 475

RESULT 11
MTRB_MYCPA
ID MTRB_MYCPA STANDARD; PRT; 565 AA.
AC Q93CE7;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Sensor histidine kinase mtrB (EC 2.7.3.-).
GN MTRB.
OS Mycobacterium paratuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1770;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 19698;
RA Urbanic K.W., Mutharia L.M.;
RT Identification and initial characterization of the mtrAB two-
RT component signal transduction system of Mycobacterium avium subspecies
RT paratuberculosis."
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Member of the two-component regulatory system mtrA/mtrB.
CC Seems to function as a membrane-associated protein kinase that
CC phosphorylates mtrA in response to environmental signals (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -!- SIMILARITY: Contains 1 HAMP domain.
CC -!- SIMILARITY: Contains 1 histidine kinase domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF410884; AAL10208.1; -
CC InterPro; IPR003594; ATPbind_ATPase.
CC InterPro; IPR004358; Actbind_pr_C.
CC InterPro; IPR003660; HAMP.
CC InterPro; IPR003661; His_kinA_N.
CC InterPro; IPR005467; His_kinase.
CC Pfam; PF00672; HAMP; 2.
CC Pfam; PF02518; HATPase_c; 1.
CC Pfam; PF00512; HsKa; 1.
CC PRINTS; PR00344; BCTRLSENSOR.
CC SMART; SM00304; HAMP; 1.
CC SMART; SM00387; HATPase_c; 1.
CC SMART; SM00388; HsKa; 1.
CC PROSITE; PS50885; HAMP; 1.
CC PROSITE; PS50109; HIS_KIN; 1.
CC Sensory transduction; Transferase; Kinase; Transmembrane;
KW Phosphorylation.
FT TRANSMEM 42 62 POTENTIAL.
FT TRANSMEM 213 287 POTENTIAL.
FT DOMAIN 235 287 HAMP.
FT DOMAIN 302 519 HISTIDINE KINASE.
FT MOD_RES 305 305 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
SQ SEQUENCE 565 AA; 61072 MW; 6D10AA571A4B50A7 CRC64;

Query Match 14.6%; Score 335; DB 1; Length 565;
Best Local Similarity 22.2%; Pred. No. 2.4e-11;
Matches 108; Conservative 111; Mismatches 221; Indels 46; Gaps 11;
```

```
QY 3 KRKLNNMIIVTMTFTVFCLIIIFFLKDTLHNSELD-----DAERSSSDINN 54
Db 35 RRSIQLRVVALTGLSLAVI---LALGFVLTQVTRVLDVKVKAATEQIERARTVGG 90
QY 55 LFHSPKPKVDI-SALDL-----NASLGNFOEIIIIYDEHNKLPETSN----- 94
Db 91 IVNGEARSLDSLSQLARNTLTSKTSASGAGTAGTDAVLVMPDGDGPRAATTAGPVDQV 150
QY 95 DNTVR--VEPGYEHRYFDRVIKKRYKIEYLIIKEPITTDQFKGYSLLIHSLNVDNVK 152
Db 151 PASLRGFKACQASYQYATVHTDPSG-PALIVSPASSQVANLEYLIIFPKNEQATIQ 209
QY 153 SLYIIALAFGVIAITITATISYVFTSTQTKPLVSLSNKMIERRDGFQNKQLNTNBEI 212
Db 210 LVRGTMITGGAVLLVLAGIALLVSRVVVPVRSAS-RIARFAEGHLSERPVRGEDDM 268
QY 213 DNLANTFNEM-----MSQIEESFNQORQVDEASHLRTPLQIIQHLNLIORWKKD 265
Db 269 ARLAMSFNDMAESLSRQITQLEEFNQLQRRFTSDVSHLRTPLTTVRMAADLIYDHSADL 328
QY 266 PAVLEESLNIISIEEMNRIIKLVEELLETKGDVNDISSEAOVTVINDEIRIHSCLKLH 325
Db 329 DPTLARSTELMVNEDRFESLNDLLEISRHDAGVAELSVEAVDURSTVQSALSNVGHLA 388
QY 326 PD--YQFDTDLTLSKNLEIKMKPHQEQFLFIFIDNAIKYDVNKKIKVTKLKNKQKII 383
Db 389 EDAGIELQVELPAEAEVIAEVDTRRVERILRNLIANAIDH-AEHKPKVKIRMAAEDTVAVT 447
QY 384 ITDHGIGIPEDQDPIFDRFVYDKSRSGGNGLSIAQKIQLNGSGSIKIKSEIN 443
Db 448 VRDYGVLGRPGEEKLVFSRFRADPSRVRRSGGTGLGLAISIEDARLHQGLEAWGEPGV 507
QY 444 GTTFKI 449
Db 508 GSCFRL 513

RESULT 12
BARA_SHIFL
ID BARA_SHIFL STANDARD; PRT; 918 AA.
AC PS9342;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Sensor protein barA (EC 2.7.3.-).
GN BARA OR SP2799 OR S2993.
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=623;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=301 / Serotype 2a;
RX MEDLINE=2272406; PubMed=12384590;
RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
RA Yu J.;
RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
RT through comparison with genomes of Escherichia coli K12 and O157.";
RL Nucleic Acids Res. 30:4432-4441 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=2457T / ATCC 700930 / Serotype 2a;
RX MEDLINE=22590274; PubMed=12704152;
RA Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,
RA Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,
RA Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,
RA Schwartz D.C., Blattner F.R.;
RT "Complete genome sequence and comparative genomics of Shigella
RT flexneri serotype 2a strain 2457T.";
```


RN
 RP SEQUENCE FROM N.A.
 RC STRAIN=O6:H1 / CFT073 / ATCC 700928;
 RX MEDLINE=22388234; PubMed=12471157;
 RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
 RA Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
 RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
 RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
 RT "Extensive mosaic structure revealed by the complete genome sequence
 of uropathogenic *Escherichia coli*.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).
 [4]
 RN SEQUENCE FROM N.A.
 RP STRAIN=O157:H7 / EDL933 / ATCC 700927;
 RC MEDLINE=21074935; PubMed=11206551;
 RX Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
 RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
 RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
 RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.F., Potamousis K.,
 RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
 RA Welch R.A., Blattner F.R.;
 RT "Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7.";
 RL Nature 409:529-533(2001).
 [5]
 RN SEQUENCE FROM N.A.
 RP STRAIN=O157:H7 / RIMD 0509952;
 RC MEDLINE=21156231; PubMed=11258796;
 RX Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
 RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
 RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
 RA Kihara S., Shiba T., Hattori M., Shinagawa H.;
 RA "Complete genome sequence of enterohemorrhagic *Escherichia coli*
 O157:H7 and genomic comparison with a laboratory strain K-12.";
 RL DNA Res. 8:11-22(2001).
 [6]
 RN SEQUENCE OF 106-918 FROM N.A.
 RP STRAIN=K12;
 RC MEDLINE=97349980; PubMed=9205837;
 RX Yamamoto Y., Aiba H., Baba T., Hayashi K., Inada T., Isono K.,
 RA Itoh T., Kimura S., Kitagawa M., Makino K., Miki T., Mitsuhashi N.,
 RA Mizobuchi K., Mori H., Nakade S., Nakamura Y., Nashimoto H.,
 RA Oshima T., Oyama S., Saito N., Sempel G., Satoh Y., Sivasubraman S.,
 RA Tagami H., Takahashi H., Takeda J., Takemoto K., Uehara K., Wada C.,
 RA Yamagata S., Horiuchi T.;
 RT "Construction of a contiguous 874-kb sequence of the *Escherichia coli*
 K-12 genome corresponding to 50.0-68.8 min on the linkage map and
 analysis of its sequence features.";
 RL DNA Res. 4:91-113(1997).
 [7]
 RN CHARACTERIZATION.
 RC STRAIN=K12 / MC4100;
 RX MEDLINE=20576358; PubMed=11022030;
 RA Pernestig A.-K., Meleforts O., Georgellis D.;
 RT "Identification of UvrY as the cognate response regulator for the BarA
 sensor kinase in *Escherichia coli*.";
 RL J. Biol. Chem. 276:225-231(2001).
 [8]
 RN CHARACTERIZATION.
 RC STRAIN=K12 / MC4100;
 RX MEDLINE=22181472; PubMed=12193630;
 RA Suzuki K., Wang X., Weillbacher T., Pernestig A.-K., Meleforts O.,
 RA Georgellis D., Babiszke P., Romeo T.;
 RT "Regulatory circuitry of the CsrA/CsrB and BarA/UvrY systems of
Escherichia coli.";
 RL J. Bacteriol. 184:5130-5140(2002).
 [9]
 RN CHARACTERIZATION.
 RP MEDLINE=2242312; PubMed=12533459;
 RX Pernestig A.-K., Georgellis D., Romeo T., Suzuki K., Tomenius H.,
 RA Normark S., Meleforts O.;
 RT "The *Escherichia coli* BarA-UvrY two-component system is needed for
 efficient switching between glycolytic and gluconeogenic carbon
 sources."

RL J. Bacteriol. 185:843-853(2003).
 CC -!- FUNCTION: Member of the two-component regulatory system uvrY/barA
 involved in the regulation of carbon metabolism via the csrA/csrB
 regulatory system. Phosphorylates uvrY, probably via a four-step
 phosphorylation.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
 (Probable).
 CC -!- PTM: Activation requires a sequential transfer of a phosphate
 group from a His in the primary transmitter domain, to an Asp in
 the receiver domain and to a His in the secondary transmitter
 domain (By similarity).
 CC -!- SIMILARITY: Contains 1 HAMP domain.
 CC -!- SIMILARITY: Contains 1 histidine kinase domain.
 CC -!- SIMILARITY: Contains 1 HPT domain.
 CC -!- SIMILARITY: Contains 1 response regulatory domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>
 or send an email to license@ebi.ac.uk).
 CC -----
 DR EMBL; D10888; BAA01710.1; -.
 DR EMBL; AE000362; AAC75828.1; -.
 DR EMBL; U29580; AAA69296.1; -.
 DR EMBL; AE016765; AAN81797.1; -.
 DR EMBL; AE005506; AAG57899.1; -.
 DR EMBL; AP002562; BAB37069.1; -.
 DR EMBL; D90894; BAA16571.1; -.
 DR PIR; F91084; F91084.
 DR PIR; S20550; S20550.
 DR HSSP; P02933; 1BXD.
 DR EcoGene; EG11367; barA.
 DR InterPro; IPR003594; ATPbind_ATPase.
 DR InterPro; IPR004358; Bact_sens_pr_C.
 DR InterPro; IPR003660; HAMP.
 DR InterPro; IPR003661; His_kinA_N.
 DR InterPro; IPR005467; His_kinase.
 DR InterPro; IPR008207; Hpt.
 DR InterPro; IPR001789; Response_reg.
 DR Pfam; PF00672; HAMP; 1.
 DR Pfam; PF02518; HATPase_C; 1.
 DR Pfam; PF00512; HiskA; 1.
 DR Pfam; PF01627; Hpt; 1.
 DR Pfam; PF00072; response_reg; 1.
 DR PIRSF; PIRSF005259; Hydr_HK_barA; 1.
 DR PRINTS; PR00344; BCTRLSENSOR.
 DR ProDom; PD000039; Response_reg; 1.
 DR SMART; SM00304; HAMP; 1.
 DR SMART; SM00387; HATPase_C; 1.
 DR SMART; SM00388; HiskA; 1.
 DR SMART; SM00073; HPT; 1.
 DR SMART; SM00448; REC; 1.
 DR PROSITE; PS00885; HAMP; 1.
 DR PROSITE; PS0109; HIS_KIN; 1.
 DR PROSITE; PS0894; HPT; 1.
 DR PROSITE; PS0110; RESPONSE_REGULATORY; 1.
 DR Sensory transduction; Transferrase; Kinase; Phosphorylation;
 Transmembrane; Inner membrane; Transcription regulation;
 Complete proteome.
 FT DOMAIN 1 9
 FT TRANSMEM 10 31
 FT DOMAIN 32 176
 FT TRANSMEM 177 196
 FT DOMAIN 197 918
 FT DOMAIN 200 252
 FT DOMAIN 299 520
 FT DOMAIN 669 785
 FT DOMAIN 822 918
 FT MOD_RES 302 302
 FT MOD_RES 718 718
 FT CYTOPLASMIC (POTENTIAL).
 FT PERIPLASMIC (POTENTIAL).
 FT POTENTIAL.
 FT CYTOPLASMIC (POTENTIAL).
 FT HAMP.
 FT HISTIDINE KINASE.
 FT RESPONSE REGULATORY.
 FT HPT.
 FT PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 FT PHOSPHORYLATION (BY SIMILARITY).

```
FT MOD RES 861 861 PHOSPHORYLATION (BY SIMILARITY).
SQ SEQUENCE 918 AA; 102453 MW; EA31D6D732F023CD CRC64;

Query Match 14.4%; Score 330.5; DB 1; Length 918;
Best Local Similarity 24.5%; Pred. No. 7.3e-11;
Matches 140; Conservative 77; Mismatches 171; Indels 183; Gaps 24;

QY 1 MTRKLRNNMIIIVTMTITFTVIFLCLIIIFLKLTLHN--SELDAAERS----- 48
DB 1 MTNYSLARWMI---LILAPVLGLLSLFFVHRVNDLQREQEDAGALIEPLAVSTE 57
QY 49 -----SSDINNLF-----HSPKVDISALDNLASLGNFQEIYIYDEHNKLFETSND 95
DB 58 YGMSLQNRISIGQLISVLHRRHSIDVAIS-----VYDE-NNRILFVTSN- 100
QY 96 NTVRVEFGY----BHRFVDRVKKRYGIEYLIIEKI-----TTQDFK 135
DB 101 --FHLDPSSMOLGNSVFPFQLTVTRDG-DIMILRTPIISESYSPDSSDAKNSQNM 157
QY 136 GY-----SLIHSLENDIVKSLYIIALAFGVATITATISVVFSTQITKPLVLSNK 190
DB 158 GYIALEDLKSRLQCKEYFISVWMLFCIG-----IALIFGRMLMRDVTGPIRNMVNT 212
QY 191 MIEIRDFQKQLNTNIE-----EIDNLANTFNEM-----MSQIE 227
DB 213 VDIRIRRG-----QLDSRVEFGMLGELDMKNGINSVMSLSAAVHEEQHNIDQATSDLR 266
QY 228 ESFNQ-----ORQFVEDASHELTPQLIIQ----- 253
DB 267 ETLQMEIIONVELDLAKRAQEAARIKSEFLANNSHELRTPLNGVIGFTRLTLTKELTPT 326
QY 254 --HLNLIQKWKDPAVLEESLNISIEEMNRKIILVEELLELTKGVDNDISSEAFQVHI 310
DB 327 QRDHLNTER-----SANNLAIINDVDFSKLEAGKILDES---I 364
QY 311 NDEIRSRHSLKQL--HPDYQFDLTSKNLEIKV-----KPHQFQQLFLIFIDNAI 360
DB 365 PFLPRLSTLDEVVTLAHSKDKGLEA--LNKSDVPDNDVIGDPLRLQIITNLVGNAI 421
QY 361 KY-DVKKKIKVTRKKNKQKI---IEITHGIGIPEEDQDFIPDRYRVKDKSRSGGG 416
DB 422 KFTENGINDILVEKRALSNKTVQIEVQIRDTGIGIPERDQSLFQAFQADASTSRHGG 481
QY 417 NGLGLSIAQKIQLNGSGSIKSEIKNGTTF 447
DB 482 TGLGLVITQKLVNEMGDISFHSQPNRGSTF 512

RESULT 14
MTRB_MYCBO STANDARD; PRT; 567 AA.
AC P59963;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DE Sensor histidine kinase mtrB (EC 2.7.3.-).
GN MTRB OR WB3273C.
OS Mycobacterium bovis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1765;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AF2122/97;
RX MEDLINE=22709107; PubMed=12788972;
RA Garnier T., Biglieri K., Camus J.-C., Medina N., Mansoor H.,
RA Pryor M., Duthoy S., Grondin S., Lacroix C., Monsemp C., Simon S.,
RA Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
RA Parkhill J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
RT "The complete genome sequence of Mycobacterium bovis.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).
CC -1- FUNCTION: Member of the two-component regulatory system mtrA/mtrB.
CC Seems to function as a membrane-associated protein kinase that
```

```
phosphorylates mtrA in response to environmental signals (By
similarity).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- SIMILARITY: Contains 1 HAMP domain.
CC -1- SIMILARITY: Contains 1 histidine kinase domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC ENBL; BX248345; CAD95365.1; -.
CC PROSITE; PSS0885; HAMP; 1.
CC PROSITE; PSS0109; HIS_KIN; 1.
CC Sensory transduction; Transferase; Kinase; Phosphorylation;
KW Transmembrane; Complete proteome.
FT TRANSMEM 42 62 POTENTIAL.
FT TRANSMEM 213 233 POTENTIAL.
FT DOMAIN 235 287 HAMP.
FT DOMAIN 302 519 HISTIDINE KINASE.
FT MOD RES 305 305 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
SQ SEQUENCE 567 AA; 61618 MW; B4C73DF5D69D064E CRC64;

Query Match 14.3%; Score 328; DB 1; Length 567;
Best Local Similarity 22.2%; Pred. No. 5.7e-11;
Matches 111; Conservative 112; Mismatches 204; Indels 72; Gaps 16;

QY 3 KRKLNNMIIIVTMTITFTVIFLCLIIIFLKLTLHN-----SELDAAERSSSDINN 54
DB 35 RRSRLQRVALTUGLSLAVI---LALGFVLTQVTRNVLDIKVRAIDQIERARTVSG 90
QY 55 LPHSKPVKDI-SALDL-----NASI-GNFQEIYIYDEHNKLFETSNDNT 97
DB 91 IVNGEETSLSDSLQLARNTLTSTDPASGAGLAGAFDAVLMYPGDPRAASTAGPVDQV 150
QY 98 VRPEPYEHRYFDRVTKRYKGYEYLIIEKPTTQPKGYSLLI---HSLNENYDNIVKSLY 155
DB 151 PNALRGF-----VKAGQAAVQYATVQ---TEGFGPALIIGTPTLSRVANL--ELY 196
QY 156 I-----IALAFGVAT-----IITATISVVFSTQITKPLVLSNKNMIEIRRDGF 199
DB 197 LIPLASEQATITLVKGTWATGGLVLLVLAGIALLVSRQVVPVRSAS-RIARFAEGH 255
QY 200 QNKQLNTNIEIDNLANTFNEM-----MSQIEESFNQQRQFVEDASHELTPQLIIQ 252
DB 256 LSERMPVRGDDMARLAVSFNDMAESLSRQIAQLAEFGNLRRTSDVSHELATPLTTVR 315
QY 253 GHNLILQKWKDPAVLEESLNISIEEMNRKIILVEELLELTKGVDNDISSEAFQVHI 312
DB 316 MAADLIYDHSADLDPTLRSTELMWSELDRFETFLNLDLLEISRHDAGVABLSVEAVDLRT 375
QY 313 EIRSRHSLKQLHPDYQFD--TDLTSKNLEIKMKPHQFQQLFLIFIDNAIKYDKNKKIK 370
DB 376 TVNNAALNGVHLAEAGIELLVDPAEQVIAEDARVERKILNLTANAIDH-AEHKPV 434
QY 371 VKTRLNKKQKIIEITHGIGIPEEDQDFIPDRYRVKDKSRSGGGNGLSLAQIKIQL 430
DB 435 IRMADEDTVAVTVRDYGVGLRGEGEKLVSFRWRSRSDPSRVRSRGGTGLGLAISVEDARL 494
QY 431 NGGSIKIKSEIKNGTTFKI 449
DB 495 HQGRLEAWGEPGEGACFRL 513

RESULT 15
MTRB_MYCTU STANDARD; PRT; 567 AA.
AC Q50496; O05890;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
```

10-OCT-2003 (Rel. 42, Last annotation update)
 MTRB OR RV3245C OR MT3343 OR MTCY20B11.20C.
 Mycobacterium tuberculosis.
 Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 NCBI_TaxID=1773;
 (1)
 SEQUENCE FROM N.A.
 STRAIN=H37RV;
 MEDLINE=96236050; PubMed=8655513;
 Via L.E., Curcio R., Mudd M.H., Dhandaythapani S., Ulmer R.J.,
 Deretic V.;
 "Elements of signal transduction in Mycobacterium tuberculosis: in
 vitro phosphorylation and in vivo expression of the response regulator
 MtrA";
 J. Bacteriol. 178:3314-3321(1996).
 (2)
 SEQUENCE FROM N.A.
 STRAIN=H37RV;
 MEDLINE=98295987; PubMed=9634230;
 Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekaiia F.,
 Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
 Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
 Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
 Sulten J.E., Taylor K., Whitehead S., Barrell B.G.;
 "Deciphering the biology of Mycobacterium tuberculosis from the
 complete genome sequence";
 Nature 393:537-544(1998).
 (3)
 SEQUENCE FROM N.A.
 STRAIN=CDC 1551 / Oshkoeh;
 MEDLINE=22208494; PubMed=12218036;
 Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 Peterson J., DeBoy R., Dodson R., Gwinn M., Haft D., Hickey E.,
 Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M., Salzberg S.L.,
 Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
 Bishai W., Jacobs W.R. Jr., Venter J.C., Fraser C.M.;
 "Whole-genome comparison of Mycobacterium tuberculosis clinical and
 laboratory strains";
 J. Bacteriol. 184:5479-5490(2002).
 -1- FUNCTION: Member of the two-component regulatory system mtrA/mtrB.
 Seems to function as a membrane-associated protein kinase that
 phosphorylates mtrA in response to environmental signals.
 -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 -1- SIMILARITY: Contains 1 HAMP domain.
 -1- SIMILARITY: Contains 1 histidine kinase domain.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 or send an email to license@isb-sib.ch)

 EMBL; U14909; AAB07807.1; -;
 EMBL; Z95121; CAB08346.1; -;
 EMBL; AE007145; AAK47685.1; -;
 TIGR; MT3343; -;
 TruncList; RV3245C; -;
 InterPro; IPR003594; ATPbind_ATPase.
 InterPro; IPR004358; Bact_sens_pr_C.
 InterPro; IPR003660; HAMP.
 InterPro; IPR003661; His_kinA_N.
 InterPro; IPR005467; His_kinase.
 Pfam; PF00672; HAMP; 2.
 Pfam; PF02518; HATPase_C; 1.
 Pfam; PF00512; HisKA; 1.
 PRINTS; PR00344; BCTRLSENSOR.

SMART; SM00304; HAMP; 1.
 SMART; SM00387; HATPase_c; 1.
 SMART; SM00388; HisKA; 1.
 PROSITE; PS00885; HAMP; 1.
 PROSITE; PS50109; HIS_KIN; 1.
 Sensory transduction; Transferase; Kinase; Phosphorylation;
 Transmembrane; Complete proteome.
 TRANSMEM 42 62 POTENTIAL.
 FT TRANSMEM 213 233 POTENTIAL.
 FT DOMAIN 235 287 HAMP.
 FT DOMAIN 302 519 HISTIDINE KINASE.
 FT MOD_RES 305 305 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 FT CONFLICT 308 308 R -> G (IN REF. 1).
 FT CONFLICT 517 517 M -> L (IN REF. 2).
 SQ SEQUENCE 567 AA; 61636 MW; B4C73C330DF1764E CRC64;
 Query Match 14.3%; Score 328; DB 1; Length 567;
 Best Local Similarity 22.2%; Pred. No. 5.7e-11;
 Matches 111; Conservative 112; Mismatches 204; Indels 72; Gaps 16;
 Qy 3 KRKLNNWIIITVTITFLCLIIIFFLKDTLHN-----SELDDAERSSSDINN 54
 Db 35 RRSQLQVWALTGLSLAVI-----LALGFVLTSQVTRVLDIKVRAAIDQIERATTVSG 90
 Qy 55 LFHSKPKVDI-SALDL-----NASL-GNFQEIYDEHNKLPETSN--DNT 97
 Db 91 IVNGEETRLSDSSQLARNTLTSKTPASGAGLAGAFDVLWVPGDGPRAASTAGPVDQV 150
 Qy 98 VRVPEGYEHRYFDRVKKRYGIEYLIIKEPIITQDFKGYSLLI--HSLNVDNIVKSLY 155
 Db 151 PNALRGF-----VRAGAAQYATVQ----TEGFSGPALIIIGTFLSRVANL--ELY 196
 Qy 156 I-----IALAFGVIAT-----IITATISYVFSTQITKPLVLSLNKMIERRDGF 199
 Db 197 LIPPLASEQATITLVRGTMATGGLVLLVLLAGIALIVSRQVVVVRAS--RIAEFAGEH 255
 Qy 200 QNKLQLNTNVEEDNLANTFNM-----MSQIEESFNQORQFVEDASHELRLPLQIIQ 252
 Db 256 LSERMPVRGEDDMARLAVSFNDMAESLSRQIAQEFGNLQRRFTSDVSHELRLPLTVR 315
 Qy 253 GHLNLIORWGKKDPAVLEESLNTSIEWNRIKLVELELTKGDVNDISSEAOQTVHND 312
 Db 316 MAADLIYDHSADLDPTLRRSTELMVSELDRFETLLNDLLEISRHDAGVAELSVAVDLRT 375
 Qy 313 EIRSRIHSLKQLHPDYQFD--TDLTSKNLEIKMKPHQFOLFLIFIDNAIKYDVKNKKIK 370
 Db 376 TVNNALGNVGHLEAEAGIELLVLPARQVIAEVDAARVERILRLNLINADH-AEHKPV 434
 Qy 371 VKTRLKNKQKIIETDHGIGIPEEDQDFIFDRFYRVDKSRSRSGGNGLSIAQKIQL 430
 Db 435 IRMAADEDTVAVTVRDYGVGLRPGEEKLVFSRFRSDPSRVRSSGGTGLGLAISVEDARL 494
 Qy 431 NGGSIKIKSEINKGTTTKI 449
 Db 495 HQGRLEAWGEPGEGACFRL 513
 Search completed: October 4, 2004, 18:03:49
 Job time : 25.194 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 4, 2004, 18:00:48 ; Search time 35.003 Seconds
(without alignments)
1239.391 Million cell updates/sec

Title: US-09-006-627-2
Perfect score: 2291
Sequence: 1 MTKRKLNNWIIVTMTITFV.....GGSIKIKSEINKGTTFKIIF 451

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR_78:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2291	100.0	451	2 F89918	hypothetical prote
2	710.5	31.0	483	2 A71609	two-component sens
3	703.5	30.7	483	2 A81247	two-component sens
4	606.5	26.5	491	2 A86824	sensor protein kin
5	481.5	21.0	459	2 D36896	two-component sens
6	471	20.6	454	2 D97239	membrane-associate
7	464.5	20.3	482	2 D86750	sensor protein kin
8	452.5	19.8	441	2 A81955	two-component sens
9	452.5	19.8	472	2 A81955	hypothetical prote
10	422.5	18.4	588	2 C89928	two-component sens
11	413.5	18.0	591	2 A71762	sensor histidine k
12	405.5	17.7	566	2 A75481	two-component sens
13	405.5	17.7	591	2 F70089	two-component sens
14	400	17.5	591	2 A81387	two-component sens
15	399.5	17.4	462	2 A83893	two-component sens
16	396.5	17.3	449	2 C95142	sensory box sensor
17	395.5	17.3	449	2 A99010	histidine kinase (
18	395	17.2	608	2 C89760	two-component sens
19	391	17.1	499	2 E96969	histidine kinase (
20	388	16.9	554	2 A89953	alkaline phosphata
21	386	16.8	459	2 A71397	two-component sens
22	386	16.8	459	2 A71397	two-component sens
23	385.5	16.8	589	2 H69691	two-component sens
24	383	16.7	473	2 H69762	two-component sens
25	383	16.7	473	2 G97085	sensor transducti
26	381.5	16.7	566	2 H97109	sensor histidine
27	380.5	16.6	610	2 A11110	two-component sens
28	379.5	16.6	610	2 A81472	two-component sens
29	376.5	16.4	596	2 A81318	two-component sens

30	370.5	16.2	475	2 H97295	membrane associate
31	370	16.2	607	2 B84153	two-component sens
32	368.5	16.1	473	2 G83932	two-component sens
33	368.5	16.1	579	2 A27650	phosphate response
34	368	16.1	498	2 G96938	sensory transducti
35	367	16.0	683	2 D97011	sensory transducti
36	365	15.9	596	2 A71690	two-component sens
37	364	15.9	479	2 AF1625	two-component sens
38	360	15.7	479	2 AD1263	two-component sens
39	359.5	15.7	480	2 D86674	sensor protein kin
40	358.5	15.6	456	2 B86834	sensor protein kin
41	358	15.6	444	2 S49545	histidine kinase -
42	358	15.6	444	2 G95092	sensor histidine k
43	358	15.6	444	2 D97960	sensor protein Cia
44	358	15.6	489	2 F72324	sensor histidine k
45	353.5	15.4	482	2 C85555	probable 2-compone

ALIGNMENTS

RESULT 1

F89918

hypothetical protein arlS [imported] - Staphylococcus aureus (strain N315)

C:Species: Staphylococcus aureus

C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 02-Aug-2002

C:Accession: F89918

R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Ogue

ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;

C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.

Lancet 357, 1225-1240, 2001

A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.

A:Reference number: A89758; MUID:21311952; PMID:11418146

A:Accession: F89918

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-451 <KUR>

A:Cross-references: GB:BA000018; PID:gl3701211; PIDN:BA842506.1; GSPDB:GN00149

A:Experimental source: strain N315

C:Genetics:

A:Gene: arlS

C:Superfamily: envZ protein; sensor histidine kinase homology

Query Match 100.0%; Score 2291; DB 2; Length 451;
Best Local Similarity 100.0%; Pred. No. 4.8e-115;
Matches 451; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MTKRKLNNWIIVTMTITFVTIFLFCIIIIFFFLKDTLHNSLDDAERSSDINNLFHSPK	60
Db	1	MTKRKLNNWIIVTMTITFVTIFLFCIIIIFFFLKDTLHNSLDDAERSSDINNLFHSPK	60
Qy	61	VKDLSALDLNASLGNFOEIIIIYDEHNNKLPETSDNDTVRVEPGVEHYRFORVKKRYGI	120
Db	61	VKDLSALDLNASLGNFOEIIIIYDEHNNKLPETSDNDTVRVEPGVEHYRFORVKKRYGI	120
Qy	121	EYLIIEKEPITTDQPKGYSLIHSLENDIVKSLIYIIFLAFGVIATITATISYVFSTQI	180
Db	121	EYLIIEKEPITTDQPKGYSLIHSLENDIVKSLIYIIFLAFGVIATITATISYVFSTQI	180
Qy	181	TKPLVLSLNKMIERRDGFQNKQLNTNYYEIDNLANTFNEMMSQIESFNQQRQFVEDA	240
Db	181	TKPLVLSLNKMIERRDGFQNKQLNTNYYEIDNLANTFNEMMSQIESFNQQRQFVEDA	240
Qy	241	SHELRTPLOIIQGHNLNIOQWKKOPAVLESNISEEMNRIIKLVBELELTKGDVND	300
Db	241	SHELRTPLOIIQGHNLNIOQWKKOPAVLESNISEEMNRIIKLVBELELTKGDVND	300
Qy	301	ISSAQTVHINDEIRSRHSLSKQLHPDYQFDTDLTSTKNLEIKMKPHQFEQLFLIFIDNAI	360
Db	301	ISSAQTVHINDEIRSRHSLSKQLHPDYQFDTDLTSTKNLEIKMKPHQFEQLFLIFIDNAI	360
Qy	361	KYDVKNKIKVKTRLKNVKQKIIEITDHGIGIPBEDQDFIFDRFYRVDKSRSGGNGLG	420
Db	361	KYDVKNKIKVKTRLKNVKQKIIEITDHGIGIPBEDQDFIFDRFYRVDKSRSGGNGLG	420


```
Db 361 KYDVNKKIKVTRLNKQKIIETDGHGIPEDQDFIDRFVRVDKSRSGGNGLG 420
Qy 421 LSTAQKIIQNGSGIKSIKSGTKTTPKIIF 451
Db 421 LSTAQKIIQNGSGIKSIKSGTKTTPKIIF 451

RESULT 2
AF1609
two-component sensor histidine kinase lisk [imported] - Listeria innocua (strain Clip112
C:Species: Listeria innocua
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
C:Accession: AF1609
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker,
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Krefit, J.; Kuhn, M.; Kunst, F.; Kurapat, G.; Madueno, E.; Maitournam, A.; Mat
ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A.; Title: Comparative genomics of Listeria species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AF1609
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-483 <GLA>
A:Cross-references: GB:AL592022; PIDN:CAC96646.1; PID:G16413888; GSPDB:GN00178
A:Experimental source: strain Clip11262
C:Genetics:
A:Gene: lisk

Query Match 31.0%; Score 710.5; DB 2; Length 483;
Best Local Similarity 33.9%; Pred. No. 8.1e-31;
Matches 163; Conservative 102; Mismatches 169; Indels 47; Gaps 10;

Qy 1 MTKRKLNNIIIVTMTITFTVIFLFLII-----IFFLKDTLHNS 41
Db 8 LKRSLSKFWTFGASAAIFLTFFLFSYAIYQIGQMLNEEPEVKELLATTTLTNQD 67

Qy 42 LDDAERSSDINNLFHSPKVDISALDNLASNGFQIIYDEHN---NKLFTSNDNTV 98
Db 68 LDTNE-----EIKYLFNNDKT-----VNRKLDQ-QVINLYDKDGHFINKYFYSRQDIT 115

Qy 99 RVEPGYEHRYF---DRVI--KKRYGIEVLIKEPITTD---FKGYSLIHSLENYDN 149
Db 116 SID---FSQYFVSGTDKFMNKPITDQKMTAQMPIVADDDNTTVIGYQVNVPLTSYNR 172

Qy 150 IVKSLYIIALAFGVIATITATISYVFSTQITKPLVSLNKMIEIRDFGNKQLQNTN- 208
Db 173 MMDRLVVTMILLGAVALFISGMLGYLLAQNFPLNPLTLARTMNDIRKNGFQRIETKTN 232

Qy 209 YEEIDNLTNTFNMMSQIESFNQORQFVEDASHELRTPLOIIOGHNLIIQWKKDPAV 268
Db 233 RDEIGELTVVFNDDMTRIETSFQOQKQFVEDASHELRTPVQIMEGHKLTLTRGKDDPAV 292

Qy 269 LEESLNISIEEMNRIKLVLELLELTKGDVNDISSEAQTVDHINDEIRSRHSLKQLHPDY 328
Db 293 LDESLNASLTELERNKMLVQEMLDLSRAEQISQTKELQITDVTNATVEQVRNPEVMYENF 352

Qy 329 QFDTDLTSKNLEIKMKPHQEQFLIFIDNAIKYDVNKKIKVTRLNKQKIIETDGHG 388
Db 353 TFTLKEDDIDLALIQHNLHQIILIIIMDNVAVKSGDGTEDVMHVYKEQOIHDVRYDYG 412

Qy 389 IGPEEDQDFIDRFVRVDKSRSGGNGGLGSLIAQKIIQNGSGSIKSIKSGTKTTPK 448
Db 413 EGISQEBIDKIFNRFVRVDKARSREKGGNGGLGLAIAKQLVEGLVGTINAVSEPDKGTIK 472

Qy 449 I 449
Db 473 I 473

RESULT 3
AB1247
sensor protein kinase [imported] - Lactococcus lactis subsp. lactis (strain IL1403)
C:Species: Lactococcus lactis subsp. lactis
C:Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001
C:Accession: AB6824
R:Boletín, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarne, K.; Weissenbach, J.; Ehrlich
Genome Res. 11, 731-753, 2001
A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis ssp
A:Reference number: AB6625; MUID:21235186; PMID:11337471
A:Accession: AB6824
```

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-491 <STO>

A:Cross-references: GB:AE005176; PID:g12724599; PIDN:AAK05691.1; GSPDB:GN00146

A:Experimental source: strain IL1403

C:Genetics:

A:Gene: kina

Query Match

Best Local Similarity 26.5%; Score 606.5; DB 2; Length 491;

Matches 152; Conservative 100; Mismatches 187; Indels 39; Gaps 12;

3 KRKRNNWLVTTMTFTVTFELFC-----LIIFELKD--TLHNSELD-----DAE 46

20 KRSLMRWAFANTVPCFTFTLFAFLYQLTSSPFKEEQQLTKRSDMSVEEVLEKADAP 79

47 RSSSDINLFLSKPV---KDISALDNLASLGNFQBIYIYDEHNNKLFETSNDNTVRVPG 103

80 LNSSLNLYIETASKIQNGESEGSLGSIIGTRKAFYIYD-LNHKLLYSTNRHFF---G 134

104 YEHRYFDRVVKRYKGIYLIKEPI--TTQDFKGYSLIHSLENDYIVKSLYIIALAF 161

135 FQOQANNEKIRGENPGVYQKIIKSQTGVVGYLQAFYDTTTHRISNLLIVLLIL 194

162 GVATIIITATSYVSTQITPLVSLSNKMIERDGFQ--KLQNTNYSEIDNLATFF 219

195 EIVALIVLAQLGYFANYFMKPLEKLYQGMQEMANDPTNDPEPIBIQSG-DEIEELAHVY 253

220 NEMSQIESFNQORQFVEDASHELTPLQIIQHLNLIQWKKDPAPVLEESLNISTEE 279

254 NDMUKMKAYLEQNRFFSDVSHELTFLAVDDGHINLNKRWKNDPEVLESQASLDE 313

280 MNRIKLVLELELTGKGVNDISSBAQTVHINDEIRSIHSLKQLHPDYQDFTDLTSKNL 339

314 VDRMKKMLEMLARLENVDLSSEELDCDGVKVCNRALKNFQLLHDDFEIVLD---NR 369

340 EI-----KMKPHQFQFLFIIDNAIKYDVKNKKIKVTRLNKQKII-EITDHGIGIPE 393

370 LIYPTHARISENHFPQGRILLDNAKYSPPDRKEIVITVSEDEQFVTSVDKIGIGISE 429

394 EDDDFIFDRVYRVDKRSRSGGGLGSLIAQKIIQLNGGSIKIKSEINKTGTTFKIIF 451

430 EDINHLFERFRADKARNREIGTGGLGSLILARLAENTQGELEVNSELGLSTFTLKP 487

RESULT 5

E83696

two-component sensor histidine kinase BH0373 [imported] - Bacillus halodurans (strain C-

C:Species: Bacillus halodurans

C>Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001

C:Accession: E83696

R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira

Nucleic Acids Res. 28, 4317-4331, 2000

A>Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and

A:Reference number: A83650; MUID:20512582; PMID:11058132

A:Accession: E83696

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-459 <STO>

A:Cross-references: GB:AP001508; GB:BA000004; NID:g10172890; PIDN:BA04092.1; GSPDB:GN00

A:Experimental source: strain C-125

C:Genetics:

A:Gene: BH0373

Query Match

Best Local Similarity 21.0%; Score 481.5; DB 2; Length 459;

Matches 123; Conservative 109; Mismatches 184; Indels 41; Gaps 11;

10 WIIIVTMTFTVTFELFCIIIFFLKDTLHN-----SELDDAERSSSDINNLPHSK 59

14 WLVTMLLINGSIFLLFKHLIF---DNETNRVAETESVSVATNAVATGDDVSLR-- 68

60 PVKDISALDNLASLGNFQBIYIYDEHNNKLFETSNDNTVRVPGVEHYRFDVRVKKRYKG 119

Db 69 -----AYLPTNGMIR-----IITDETASVLTISSEGEMLDLPTKYQNS--OSVENVRTFDD 116

Qy 120 IEYLIIEKPTTQFKGYSLIHSLENDY---TVKSLYIIALAFGVIIATITATISYVF 176

Db 117 IPYAVSVPLIWD--GSIVMLEVTENLEDQSORTLRTLLLVLVIASVIVLIPALPLGNML 174

Qy 177 STQITKPLVSLSNKMIERDGFQNKQL--NTNVEEDNLANTNEMMSQIESFNQORQ 235

Db 175 SRLLTPIQSLLTKMKHIQETGDYEKIKLPSSKDELYTGMGTNQMIDMLEQNYKQQQ 234

Qy 236 FVEDASHELRTPLQIIQHLNLIQWKKDPAPVLEESLNISTEEMNRIIKLVELELTK 295

Db 235 FVSDASHELKTPLVIESYASMLRWKMKPELDEAVEAHEBAVRKAWTEQMLQLAN 294

Qy 296 GVDNDISSEAQTVHINDEIRSIHSLKQLHPDYQDFTDLTSKN--LEIKMKPHQFQFL 353

Db 295 EDAREVLEKKEV---DLLEVAEKACQHIHVSFERIQVVKQGEVFRVWADENKLVQVY 350

Qy 354 IFIDNAIKYDVKNKKIKVTRLNKQKIIETDHGIGIPEDDQFIIDFRVYRVDKRSRS 413

Db 351 ILLDNARKY--SQGRITVTIKETSETCSLVDQGISGDKLEKIFDRFVYRVDKARARE 408

Qy 414 QGNGGLGSLIAQKIIQLNGGSIKIKSEINKTGTTFKII 450

Db 409 TGGAGLGLSLIAKAIYDAGGGEIAIESREGKGTTLVW 445

RESULT 6

D69859

two-component sensor histidine kinase homolog ykoH - Bacillus subtilis

C:Species: Bacillus subtilis

C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 20-Jun-2000

C:Accession: D69859

R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter

A.; Ehrlich, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler

Db 59 EGVALQDMLQAY-----LPANGMVRVVGQKAVMTITKEKAYKDFPLSF 103

Qy 119 -----GIEYLIKEPITTDQFKGYSI-LHSLNENYDNIYKSLYIILALRGVIA 165

Db 104 HSGETADVRKPDGKLFABAAVPIWTDGQVVSQQLVERLENTEESFLKILIIAASAAV 163

Qy 166 TITATISYVFSTQITKPLVSLSNKMLEIRRDGFQNKQLQ-LNNYEEIDLANTFNEMMS 224

Db 164 CIASFAGSLARIINPIRLMITMKDIORDEKFTISLEGOSNDELQMGUFTFNEMAM 223

Qy 225 QIEESFNQORQFVEDASHELRTPLQIIQGHNLNLIQRMGKDPAPVLEESNISIEENRII 284

Db 224 MLKHEYDQOQFQVODASHELKTLPTIIESYSLMKRWGAKKPEVLEESIEAHSEAVHMK 283

Qy 285 KLVLEELLETQGVNDISSAQVTHINDEIRSIHSLKQLHP-DYQFDITDLTSKNLEIKM 343

Db 284 KLTNQLLALAKSH-QGLEVDLKTIDLIKAARAVMTQLSVYQRIIDLETD--KESLLVKA 340

Qy 344 KPHQFQELIFIDNAIKYDVKNKIKVTRLNKQKIIIBITDHGIGIPEEDQDFIDRF 403

Db 341 DESRIKQLLILDLNAIKY--SEKPIEMSAGTNRPFVSRBEGIGIPEEHIPHLIFERF 398

Qy 404 YRVDKSRSSQGGNGLGSLIAQKIIQLNGGSIKIKSEINKGTTFKIF 451

Db 399 YRADEARNRTGGTGLGSLIAQIADEHGIETLSVKSPGGGTAVTWF 446

RESULT 7

F97239

membrane-associated sensory histidine kinase with HAMP domain [imported] - Clostridium a

C;Species: Clostridium acetobutylicum

C;Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change: 14-Sep-2001

C;Accession: F97239

R;Nolling, J.; Breton, G.; Omeichenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee, J. Bacteriol. 183, 4823-4836, 2001

A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium C

A;Reference number: A96900; MUID:21359325; PMID:21359325

A;Accession: F97239

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-482 <KUR>

A;Cross-references: GB:AE001437; PIDN:AAK80705.1; PID:g15025798; GSPDB:GN00168

A;Experimental source: Clostridium acetobutylicum ATCC824

C;Genetics:

A;Gene: CAC2760

Query Match 20.3%; Score 464.5; DB 2; Length 482;

Best Local Similarity 31.4%; Pred. No. 1e-17;

Matches 149; Conservative 73; Mismatches 208; Indels 45; Gaps 11;

Qy 10 WIIVTMTITVFITFLCLIII--FFLKDITLHNSLDDAERSSSDINNLFHSPKVKIISAL 67

Db 13 WQLLRMFLLLICFLIILEIFQYFFLR--HYICKKEMLESIRNNVAHL-----MTTI 64

Qy 68 DLNLSLGNFQBIII-----YDEHNKLFETSNNTVVRPEGYEHRFDRVIK 115

Db 65 DSEASLKNNGYLLKSLYFDTTSSVDKNGNVLFIKDKHPTSPPIISASEYKAIMKR 124

Qy 116 RYKGIEYLIK-----EPITTDQFKGYSI-----LIHSLNENYDNIYKSLYIILALRGVIA 162

Db 125 EGNLEGYTIKNGKGEFOITWRKIGNLSPGLQIUTTPYHSAQFLYQIYIIVAC- 183

Qy 163 VIATITATISYVFSTQITKPLVSLSNKMLEIRRDGFQNKQLQNLNTYEEIDLANTFNEM 222

Db 184 VLILILSLIGGTLIKHTLKPFLYRIATLNNLTIVGDLNKLSENNOQVEIDLATAFNM 243

Qy 223 MSQIEESFNQOR-----QVVEDASHELRTPLQIIQGHNLNLIQRMGKDPAPVLEESINI 275

Db 244 FSRIEISFNREKELKXMQQFISDASHELRTPLTSIRGFVEVLIGSYKDEKLLKALNT 303

Qy 276 SIEMNRIIKVLELLELTQGVNDISSAQVTHINDEIRSIHSLKQLHPDYQFDITLT 335

Db 304 ILIESERLTELNNLLTLRLD-RSTHTMTKQNMDSIIIEIYPQLKILAKNRQVNLNLY 362

Qy 336 SKULSIKMKPHQFQELFLIFIDNAIKY-DVKNKIKVTRLNKQKIIIEITDHGIGIPEE 394

Db 363 DNNAYFLGNKQIKQIVFNIVQNSINYTDKETGIIITLSLSIEDTFLIKISONGIGIASK 422

Qy 395 QDQDFIDRFYRVDRKSRSSQGGNGLGSLIAQKIIQLNGGSIKIKSEINKGTTFKII 449

Db 423 DLPHIFDRFRSETHRSRESGGYGLSLIAQKIIQSHGGEISTKSIINQGTSTFYI 477

RESULT 8

D86750

sensor protein kinase [imported] - Lactococcus lactis subsp. lactis (strain IL1403)

C;Species: Lactococcus lactis subsp. lactis

C;Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001

C;Accession: D86750

R;Bolotin, A.; Wincker, P.; Mauger, S.; Jalllon, O.; Malarne, K.; Weissenbach, J.; Ehrlich Genome Res. 11, 731-753, 2001

A;Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis ssi

A;Reference number: A86625; MUID:21235186; PMID:11337471

A;Accession: D86750

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-441 <STO>

A;Cross-references: GB:AE005176; PID:g12723951; PIDN:AAK05102.1; GSPDB:GN00146

A;Experimental source: strain IL1403

C;Genetics:

A;Gene: kinE

Query Match 19.8%; Score 452.5; DB 2; Length 441;

Best Local Similarity 27.8%; Pred. No. 4e-17;

Matches 126; Conservative 98; Mismatches 176; Indels 53; Gaps 12;

Qy 18 TTVTIPL-FCIIIFCLIIIFLKDITLHNSLDDAERSSSDINNLFHSPKVKIDALDNLASLGNF 76

Db 16 TTVRIFLIFLIIITIVSIGIVAFSTLQIRKSEGAALLN-----SVQTAATNGQV-NW 66

Qy 77 QELIYDEHNK-----LFETSNNTVVRPEGYEHRFDRVIKRYKGYEYLIKEPI-- 129

Db 67 DEFKLDSERDEKATFVRLTSPSGORE-----ESQGTTFNFKSRKSWGNFLLRDDIFL 119

Qy 130 ----TTQDFKGYSLIHLNENYDNIYK-----LYIILALRGV--IATITATISYVFST 178

Db 120 YSKYSKNGTKAELWL-----NINVVKSMIRAILVIIAVNIFLFIALLIILQKAARTISQ 175

Qy 179 QITKPLVSLSNKMLEIRRDGFQNKQLQNLNTYEEIDLANTFNEMMSQIEESFNQORQFVE 238

Db 176 PLTDDIVVATDGDVK-----QVPVSKNPLEVHQLSQSFNRLNRLNKLQIENEQOQFVS 227

Qy 239 DASHELRTPLQIIQGHNLNLIQRMGKDPAPVLEESNISIEENRIIKVLEELLETQGV 298

Db 228 DASHELRTPLQIIQGHNLNLIQRMGKDPAPVLEESNISIEENRIIKVLEELLETQGV 298

Qy 299 NDSSAQVTHINDEIRSIHSLKQLHPDYQFDITDLTSKNLEIKMKPHQFQELFLIFID 357

Db 288 LEIKERINLSFTDQVSEIQLALRQKINFIQITDLF-----VSMKVALYKIIIFLE 342

Qy 358 NAIKYDVKNKIKVTRLNKQKIIIBITDHGIGIPEEDQDFIDRFYRVDRKSRSSQGGN 417

Db 343 NAGKYSFNSIEITNITLENHQVDFQVADGIGVPEEKDNIFERFYRIDKSRSSSIPGT 402

Qy 418 GLGLSTAQKIIQLNGGSIKIKSEINKGTTFKII 450

Db 403 GLGLALAKYQAINDAKVFNNSPKGSTFHLL 435

RESULT 9

AE1955

two-component sensor histidine kinase alr1192 [imported] - Nostoc sp. (strain PCC 7120)

C;Species: Nostoc sp. PCC 7120

A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120

C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002

C:Accession: AB1955
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Saeamoto, S.; Watanabe, A.; Iriguchi, Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S. DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena sp. PCC 7120
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AB1955
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-472 <KUR>
A:Cross-references: GB:BA000019; PIDN:BAB73149.1; PID:gl7130539; GSPDB:GN00179
A:Experimental source: strain PCC 7120
C:Genetics:
C:Superfamily: envZ protein; sensor histidine kinase homology

Query Match 19.8%; Score 452.5; DB 2; Length 472;
Best Local Similarity 26.6%; Pred. No. 4.3e-17;
Matches 129; Conservative 110; Mismatches 187; Indels 59; Gaps 13;

QY 6 LNNWMI-----IVTMTITFTVFIPLFCL-----IIFFFLKDTLHNSLDD- 44
DB 4 IKQWLNIDPFLQLRLTVGMSFALVLSLWTSWMOQILI-----DSHKNEVEQI 58
QY 45 AERSSDINNLFHSPKVDISALD--LNASLGNFOEIIYDEHNKLPFETSDNTVRVEP 102
DB 59 AKRLPQDVQ--IYSQWQPDTSLEKAIN-NFANTNTLIWFKNSNNKIL--SKTDNLDLLP 113
QY 103 GYEHYFDRVIVKRYKGYEYLIIKEPIITQDPKGVSLIHLSEYDNIIVKSLYII----- 157
DB 114 DF-----LVTKLNLNTQMPIKAQVYQVQNYFVLCGTIKIQGLKGLFVVKQITR 165
QY 158 -----ALAFGVIAITATISYVFSQIT-KPLVLSNKMIEIRRDGF-QNKLQL 205
DB 166 EOTMFVWVRSICITSLIILITVAIAFYKRSLOPLRQLSQMTAVIASADLGOAQYL 225
QY 206 NNTVEIDNLANTFNEMMSQIEESFNQORQFVEDASHELRLPQIIQGLHNLIRWGKDD 265
DB 226 DNAPEVKELAQTLTMLSRLSQSQWEQREFVSNVSHLRLPLTVHGLQSVLRQNNL 285
QY 266 PAVLEESLNISEENRILKVEELLETGKDNDISSEAQTVHNDIRSIHSLKQLH 325
DB 286 TSIQEALEIAEASEARTIRLQDLDLARADSGLYLFQMKTYVLND-IVAEIVVMAEKY 344
QY 326 PDYQPDTLTSLNKLKMKPHQFQLFIIFDINAIKYDVKNNKIKVTRLNKQKIIET 385
DB 345 SDRLITIEFTPIEVKVDYSRFKQVLLNLDNAVYSEADTPIFNLDQLQDTALIQVC 404
QY 386 DHGIGIPEDQDFIDFRYRVKRSRSGGNGLSIAQKIIQLNGGSIKIKSEINKGT 445
DB 405 DQGYIPLQHQARIFERFVRIDESRSQATGGCGLGLSVTKLVEGMSGVSQSKLGBGS 464
QY 446 TFKII 450
DB 465 IFTII 469

RESULT 10
C89928
hypochemical protein srrB [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C:Accession: C89928
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.; C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiranatsu, K. Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A:Reference number: A89758; MUID:21311952; PMID:11418146
A:Accession: C89928
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-588 <KUR>

A:Cross-references: GB:BA000018; PID:gl3701290; PIDN:BA842584.1; GSPDB:GN00149
A:Experimental source: strain N315
C:Genetics:
A:Gene: srrB

Query Match 18.4%; Score 422.5; DB 2; Length 588;
Best Local Similarity 23.7%; Pred. No. 2.3e-15;
Matches 137; Conservative 90; Mismatches 195; Indels 155; Gaps 12;

QY 11 IIVTMTITFTVFIPLFCLIIIFFLKDTLHNSLDDAERSSDINNLFHSK--PVKDIALD 68
DB 23 LIVTVILLISALITFMQYFTQET-ENAIREDARRISSLVEQ-SHNKEBAIKYSQTLI 80
QY 69 LN-----ASLGNFOEIIYDEHNKLPFETSDNTVRVEPGYEHYFDRVI 113
DB 81 ENPGGLMIINNKHQRSTASLSNKKQMLNVEVNDHDFDVFDKGKSVTRNV-----TI 133
QY 114 KRYKGYEYLIIKEPIITQ--DFKGYs--LLIHSLENDVNIIVKSLYIIAALAFGVIAIT 169
DB 134 KEKGSQTYILLGYPTKAKQNSHSKYSGVFIYDKLSIEDTNNAITIITITAVIFLIT 193
QY 170 ATISYVSTOITTKPLVLSNKMIEIRRDGFQNKQLANTNVEEIDNLANTFNEMMSQIEE- 228
DB 194 TVFAFFLSRLTKPLRLRQDQATRVSEGDYSYKPSVTYK-DEIGQLSQAFNQMSIEEH 252
QY 229 -----SFNQORQ----- 235
DB 253 VDALSTSKNIRDSLINSWVEGLGINSRQIILSNKXKANDIMDNIDEDAKAFLRLQIBDT 312
QY 236 -----FVDA 240
DB 313 FKSQTEMRLDENARFVVTTSYIDKIQGGSGVVVTVRDMTNEHNLQMKKDFIANV 372
QY 241 SHELRTPQLIIQGLHNLIRWGKDDPAVLEESLNISEENRILKVEELLETGKDND 300
DB 373 SHELRTPISLQGGTESIVDGIPTVEDEIKESLAVILDESRLNRLNELLNVARMADAG 432
QY 301 ISSEAQTVHNDIRSIHSLKQLHPDYQFDTLTSLNLEI-----KMKPHQFQLF 352
DB 433 LS-----VNKEVQPIAALLDKMKIKYRQADDLGLNMTFNKYCKRVWSYDMRMDQVL 485
QY 353 LIFIDNAIKYDVKNNKIKVTRLNKQKIIETDHGIGIPEDQDFIDFRYRVKRSR 412
DB 486 TNLIDNASRYTKPGDEIAITCDENESDILYIKGTGTGIAPEHLQQVDFREYKVDAAATR 545
QY 413 SQGGNGLSIAQKIIQLNGGSIKIKSEINKGTTFKI 449
DB 546 GKQGTGLGLFCRMIIEHGGSIDVKSELGKGTTFII 582

RESULT 11
AP1762
two-component sensor histidine kinase phoR [imported] - Listeria innocua (strain Clip112
C:Species: Listeria innocua
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
C:Accession: AP1762
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H. Science 294, 849-852, 2001
A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapat, G.; Madueno, E.; Maitournam, A.; Ma ok, C.; Schluster, T.; Simoes, N.; Tierrez, A.; Vasquez-Boland, J.A.; Voss, H.; Wehland, A:Title: Comparative genomics of Listeria species
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AP1762
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-591 <GLA>
A:Cross-references: GB:AL592022; PIDN:CAC97870.1; PID:gl6415180; GSPDB:GN00178
A:Experimental source: strain Clip11262
C:Genetics:
A:Gene: phoR

Qy	279	EMNRIIKVLELLELT--KGDVNDISSE----	AQTVHINDEIRSIHSLKOLHPDYQFDT	332
Db	286	EAERLERLVNDLLDLAQLEGESYMKDEPIA	QOLVYEVVEQFRLAAQRLQVHLDLDE	345
Qy	333	DLTSKNLEIXKPKHQFQFLIFIDNAIKYD	VKNKKIKVTRLNKQKIIETDGHGIGIP	392
Db	346	DVI-----ISGDPDRIEQVIRNLLDNAI	KYABETTISLTIGREGSRAQFILSDQGIGIP	400
Qy	393	EEDQDFIPDRFYRVDKSRSRSGNGLGLSIA	OKIILQNGSGIKIKSEINKGTTFKI	449
Db	401	EDKLALVFERFYRVNKSRRKGGTGLGLAIA	SQIIKKHGGTIDITSQELKGTVTI	457

Search completed: October 4, 2004, 18:10:23
Job time : 39.003 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 4, 2004, 18:09:40 ; Search time 120.491 Seconds
(without alignments)
1204.500 Million cell updates/sec

Title: US-09-006-627-2

Perfect score: 2291

Sequence: 1 MTRKRLRNWIIVTMTIFV.....GGSIKIKSEINKGTTPKIF 451

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1351062 seqs, 321799191 residues

Total number of hits satisfying chosen parameters: 1351062

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2291	100.0	451	12	US-10-282-122A-70332 Sequence 70332, A
2	703.5	30.7	483	12	US-10-282-122A-60936 Sequence 60936, A
3	464.5	20.3	482	12	US-10-282-122A-51971 Sequence 51971, A
4	426	18.6	585	12	US-10-282-122A-71330 Sequence 71330, A
5	425.5	18.6	579	9	US-09-815-242-5320 Sequence 5320, Ap
6	425.5	18.6	583	9	US-09-815-242-12597 Sequence 12597, A
7	424	18.5	613	12	US-10-282-122A-45658 Sequence 45658, A
8	422.5	18.4	583	9	US-09-866-292-2 Sequence 2, Appl
9	422.5	18.4	588	12	US-10-282-122A-44184 Sequence 44184, A
10	413.5	18.0	609	9	US-09-815-242-10749 Sequence 10749, A
11	412.5	18.0	612	12	US-10-282-122A-42582 Sequence 42582, A
12	408	17.8	611	12	US-10-282-122A-57929 Sequence 57929, A
13	401	17.5	589	12	US-10-282-122A-70610 Sequence 70610, A
14	397.5	17.4	608	9	US-09-925-637-70 Sequence 70, Appl
15	397.5	17.4	608	14	US-10-084-205-70 Sequence 70, Appl

Sequence 70, Appl	16	397.5	17.4	608	16	US-10-712-713-70
Sequence 315, App	17	396.5	17.3	449	16	US-10-474-776-315
Sequence 13266, A	18	395.5	17.3	449	9	US-09-815-242-13266
Sequence 13559, A	19	395.5	17.3	449	9	US-09-815-242-13559
Sequence 74002, A	20	395.5	17.3	449	12	US-10-282-122A-74002
Sequence 5677, Ap	21	395	17.2	608	9	US-09-815-242-5677
Sequence 12669, A	22	395	17.2	608	9	US-09-815-242-12669
Sequence 12999, A	23	395	17.2	608	9	US-09-815-242-12999
Sequence 44048, A	24	395	17.2	608	12	US-10-282-122A-44048
Sequence 71629, A	25	392	17.1	608	12	US-10-282-122A-71629
Sequence 45655, A	26	387.5	16.9	466	12	US-10-282-122A-45655
Sequence 74435, A	27	386.5	16.9	450	12	US-10-282-122A-74435
Sequence 46660, A	28	385.5	16.8	457	12	US-10-282-122A-46660
Sequence 72170, A	29	384	16.8	400	12	US-10-282-122A-72170
Sequence 52125, A	30	381.5	16.7	566	12	US-10-282-122A-52125
Sequence 60921, A	31	380.5	16.6	610	12	US-10-282-122A-60921
Sequence 60872, A	32	376.5	16.4	596	12	US-10-282-122A-60872
Sequence 48548, A	33	376	16.4	455	12	US-10-282-122A-48548
Sequence 70820, A	34	374.5	16.3	610	12	US-10-282-122A-70820
Sequence 52231, A	35	373	16.3	569	12	US-10-282-122A-52231
Sequence 46609, A	36	368	16.1	591	12	US-10-282-122A-46609
Sequence 50, Appl	37	363.5	15.9	436	14	US-10-169-048-50
Sequence 45578, A	38	363.5	15.9	458	12	US-10-282-122A-45578
Sequence 134, App	39	359	15.7	422	15	US-10-264-213-134
Sequence 284, App	40	358	15.6	444	16	US-10-474-776-284
Sequence 48757, A	41	355.5	15.5	443	12	US-10-282-122A-48757
Sequence 53525, A	42	353.5	15.4	429	12	US-10-282-122A-53525
Sequence 10890, A	43	351.5	15.3	521	14	US-10-156-761-10890
Sequence 43154, A	44	349.5	15.3	480	12	US-10-282-122A-43154
Sequence 46948, A	45	349.5	14.8	357	12	US-10-282-122A-46948

ALIGNMENTS

RESULT 1

US-10-282-122A-70332

; Sequence 70332, Application US/10282122A

; Publication No. US20040029129A1

; GENERAL INFORMATION:

; APPLICANT: Wang, Liangsu

; APPLICANT: Zamudio, Carlos

; APPLICANT: Malone, Cheryl

; APPLICANT: Haselbeck, Robert

; APPLICANT: Ohlsen, Kari

; APPLICANT: Zvekind, Judith

; APPLICANT: Wall, Daniel

; APPLICANT: Trawick, John

; APPLICANT: Carr, Grant

; APPLICANT: Yamamoto, Robert

; APPLICANT: Forsyth, R.

; APPLICANT: Xu, H

; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

; FILE REFERENCE: ELITRA.034A

; CURRENT APPLICATION NUMBER: US/10/282.122A

; CURRENT FILING DATE: 2003-02-20

; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/206,848

; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60/207,727

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: 60/230,335

; PRIOR FILING DATE: 2000-09-06

; PRIOR APPLICATION NUMBER: 60/230,347

; PRIOR FILING DATE: 2000-09-09

; PRIOR APPLICATION NUMBER: 60/242,578

; PRIOR FILING DATE: 2000-10-23

; PRIOR APPLICATION NUMBER: 60/253,625

; PRIOR FILING DATE: 2000-11-27

; PRIOR APPLICATION NUMBER: 60/257,931

; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: 60/267,636


```
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELTRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; PRIOR FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 51971
; LENGTH: 482
; TYPE: PRT
; ORGANISM: Clostridium acetobutylicum
US-10-282-122A-51971

Query Match          20.3%; Score 464.5; DB 12; Length 482;
Best Local Similarity 31.4%; Pred. No. 7.6e-28;
Matches 149; Conservative 73; Mismatches 208; Indels 45; Gaps 11;

Qy 10 WIIVTMTTFTVTFIFLCIIII--FPLKDTLHNSLDDAERSSDINNLFHSPKPKVDISAL 67
Db 13 WQLLRMPFLLCIFLIIIEIFQYFLR---HYICKKEMLLSRRINNVAHL-----MTTI 64
Qy 68 DLNASLGNFQIIII-----YDEHNKLPETSDNTNTRVPEGYEHRYFDRVIKK 115
Db 65 DSEASLKNGSVLYKSKSYFTTSSVDKNGNVLFLHKDKHPTSPPIISASEYKAIKMR 124
Qy 116 RYKGEYLIK---EPTTQDFKGYSL-----LIHSLNEDN----IVKSLYIIIALAFG 162
Db 125 EGNLEGYITKNCCEPOITTRKIGNLNSPLGLIQLTTPHSAQAFLYKQIYIYIVAC- 183
Qy 163 VIATITATISYVFSTQITKPLVLSNKMIEIRRDGFQNKQLQNTNVEEDNLANFNEM 222
Db 184 VLILSILIGTTLIKHTKPLRYAHTLNNLTVGDLNKLSENNOQVEIDHLATAPNNM 243
Qy 223 MSQTEESFNQR-----QFVEDASHRLTFLQIIQHLNLIORWGKKDPAVLESINI 275
Db 244 FSRIEISFNREKELKKKQKQFISDASHRLTFLTSIRGFVEVLIGSYKDEKCLKALNT 303
Qy 276 SIEENMRIKVLLELLETKGDVNDISSEAQTVHINDEIRSIHSLKQLHPDYQFDTLT 335
Db 304 ILIESERLTELNNLLTLTRLD-RSTHTMTKQNMDSIIIEIYQOLKILAKNRQNLNLY 362
Qy 336 SKNLEIKMKPHQFQLFIFIDNAIKY--DVKNKKIKVKTRLNKQKIIETDHDGIGIPEE 394
```

```
Db 363 DNNAYFLGNKNQIKQVIFNIVQNSINTYDKRETGIIITLSIEDTLFILKISDNGIGIASK 422
Qy 395 DQDFIFDRFYRVKRSRSRSGGNGLGSLIAQKIQLNGSGSIKIKSEINKGTTFKI 449
Db 423 DLPHFDRFRFRSETHRSRSGGYGLGLSLIAKSIILQSHGGEISTKSIINQGTIFYI 477

RESULT 4
US-10-282-122A-71330
; Sequence 71330, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELTRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; PRIOR FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 71330
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Staphylococcus haemolyticus
US-10-282-122A-71330

Query Match          18.6%; Score 426; DB 12; Length 585;
Best Local Similarity 23.0%; Pred. No. 1e-24;
Matches 134; Conservative 105; Mismatches 184; Indels 160; Gaps 14;

Qy 10 WIIVTMTTFTVTFIFLCIIIF---FLKDTLHNSLDDAERSSDINNLFHSPKPKVDISA 66
Db 13 WLTLFLVTIVLLLSAALITFIQYITQOTENAIRDASR-----ISHLVEQADNKL-A 67
Qy 67 LDNLASLGNFQIIIIYDEHNKLPETSDNTN--VRVPEGYEHRYFDRV----- 112
Db 68 IEHSQQLIDGGGVIIWANKSSSIKSSNSNTKDKMLEIHKNSQFQVFSQKSTTQNT 127
Qy 113 IKGYKGEIYLIKEPTTQDFKG--YS--LLIHSLENYONIVKSLYIIALAFVIATII 168
Db 128 ISNNGNSHSHYLLGYPMKAQENANSKYSAVFIYQDLKSIEDTNNAITIILITAIPIAV 187
```



```
Db 539 GMGIPKENVDKIFERYFYRVDKARSQMGGTGLGIAITAKEMI EAHGGSIAWAKSEEGKGT 597

RESULT 8
US-09-866-292-2
; Sequence 2, Application US/09866292
; Patent No. US20020061572A1
; GENERAL INFORMATION:
; APPLICANT: Wallis, Nicola G.
; TITLE OF INVENTION: Histidine Kinase
; FILE REFERENCE: GM10202
; CURRENT APPLICATION NUMBER: US/09/866,292
; CURRENT FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 09/272,414
; PRIOR FILING DATE: 1999-03-19
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 583
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
; US-09-866-292-2

Query Match 18.4%; Score 422.5; DB 9; Length 583;
Best Local Similarity 23.7%; Pred. No. 1.9e-24;
Matches 137; Conservative 90; Mismatches 195; Indels 155; Gaps 12;

Qy 11 IIVTMTITFTVFLFCLIIIFPLKOTLHNSLDDAERSDDINNLPHSK--PVKDISALD 68
Db 18 LIVTVLLLSIALITFMQYFTQET-ENAIREDARRISLVEQ-SHNKEAIAKYSQTLI 75
Qy 69 LN-----ASLGNFQEIILYDEHNKLFETSNDNTVVEPCGYEHRYEDRVI 113
Db 76 ENPGGLMIINXKQRSTASLSNKKOMLNEVNDHFDVDFDKGKSVTRNV-----TI 128
Qy 114 KRYKGYEYLIIKEPTTQ--DFKGYG--LLIHSLENYDNIVKSLYIIALAFGVIAITIT 169
Db 129 KEGSSQTYILLGYPTKQKNSHSGYGVFIYKDLKSIEDTNNAITIITITAVIFLTIT 198
Qy 170 ATISYVSTQITKPLVSLNKMIEIRRDGFQNKQLNTNVEIDNLANTFNEMMSQIEE- 228
Db 189 TVFAFLSSRITKPLRLRLDQATRVSEGDYSYKPSVTTK-DEIGQLSQAFNQMSIEEH 247
Qy 229 -----SFNQORQ----- 235
Db 248 VDALSTSKNIRDSLNSMVEGLVGINESRQIILSNKQANDIMDNIDEDAKAFLLRQIEDT 307
Qy 236 -----FVEDA 240
Db 308 FKSQKQTEMRLDLENNARFFVVTTSYIDKIEQGGSGVVVTVRDMTNEHNLDMQKKDFIANV 367
Qy 241 SHELRTPLQIQHNLNIQRWKKOPAVLBSLNISIEEMNRRIKLVLELLELTGKDND 300
Db 368 SHELRTPLISLQGYTESIVDGIPTPEDEIKESLAVLDKSLNRLNELLNVARMDAEG 427
Qy 301 ISSEAQVHNDERSIHSKQLHPDYQFDTLTSKNLEI-----KKKHPQEQLP 352
Db 428 LS-----VNKEVQPIAALLDKMKIKYRQOADDGLNMTFTNYCKRWWSYDMDRMDQVL 480
Qy 353 LIFIDNAIKYDVKNNKIKVTRLNKKQIIEITDHGIGIPEEPQDFTDFRFRVDRKSRSR 412
Db 481 TNLIDNASRTYKDEDAITCDENESIEDLYIKDTGTGIAPEHLQVQFDRFYKVDAAATR 540
Qy 413 SQQNGNLGLSIAQKIQLNGGSIKIKSEINKGTTFKI 449
Db 541 GKQGTGLGLFICKMIEEHGGSIDVKSGLKGTTFII 577

RESULT 9
US-10-282-122A-44184
; Sequence 44184, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
```

241	Qy	SHELRTPLQIIIOGHNLNLRQWKKDPAPVLESLSNISIEBNNRIIKULVEBELLELTGCDVND	300
242		: : : : : : : : : : : : : : : : : : : :	
373	Db	SHELRTPLSLLQGYTESIVDGLVTEPDEIKESLAVLDESRLNRLVNELLANVARMDAEG	432
374		: : : : : : : : : : : : : : : : : : : :	
301	Qy	ISSEAQTVHINDEIRSRTHSLKQLHPDQPPDTLTSKNLEI-----KWKPHQEQLP	352
302		: : : : : : : : : : : : : : : : : : : :	
433	Db	LS-----VNKEVQPIAALLDKWKIKYRQADDGLNNMTFNCKKRWWSYDMDRMDQVL	485
434		: : : : : : : : : : : : : : : : : : : :	
353	Qy	LIFIDNAIKYDVNKNKKIKVTRLNKNQKIIETDHHGIGIPEEQDPIFDRFYRVKRSR	412
354		: : : : : : : : : : : : : : : : : : : :	
486	Db	TNLLIDNASRYTKPGDEIAITCDENESEDILYIKDTGTGIAPEHLQOVDFRYKVDAAFR	545
487		: : : : : : : : : : : : : : : : : : : :	
413	Qy	SGGNGLGLSLAQKIIQLNGGSIKIKSEINKGTTFKI	449
414		: : : : : : : : : : : : : : : : : : : :	
546	Db	GKQGTGLGLFICKMLIEBHGGSIDVKSELGKGTTFII	592
547		: : : : : : : : : : : : : : : : : : : :	

RESULT 10

```

US-09-815-242-10749
; Sequence 10749, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 10749
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-815-242-10749

```

RESULT 11

```

US-10-282-122A-42582
; Sequence 42582, Application US/10282122A
; Publication No. US20040029129A1
GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELIPRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1

```



```
; SEQ ID NO 42582
; LENGTH: 612
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-10-282-122A-42582

Query Match
Best Local Similarity 18.0%; Score 412.5; DB 12; Length 612;
Matches 124; Conservative 119; Mismatches 185; Indels 157; Gaps 13;

Qy 17 ITFTVFIFLCIIII--FPLKDTLNSLDDAERSSDI-----NNLFHSPKVPKDISALDNLN 70
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 17 LSFILLLLIAIQTIGFYFIRELEATTISDFKKNMDSQVQLSNTLSTQMKNKDLERSD 76
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 71 ASL-----GNFQIIIIYDEHNKLFETSDNTVVRPGYEHRYFD--RVIKKRYKG 119
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 77 ANLKALSDFSNADISEARIVDDKG--IIRATNDLNQONIGKKNYDRDLNDFTSKRYQA 134
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 120 IE-----YLIKEPI---TTQDFKGYSLIHSLEN--YDNIVKSLYIILAFGVITATIITA 170
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 135 LDNDKRVYVNV-QPIQSTGETVIGLVYKSNLENKYQEINTASIFFTA-SILAAISI 192
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 171 TISVFTQTITKPLVLSNKNMIEIRRDGFQNKQLNTNYBEIDNLANTFNMMSQIEB-- 228
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 193 IVTLIARSITKPIGEMREQAIRIARGDYAGKVEH-GKDELQGLAETFNQLSERIEAQ 251
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 229 ----- 228
Db 252 ETWEARNRLDSVLTHMTDGVATDRRGKVTINEMALSLLNVKNENVICTSILLELDIE 311
Qy 229 ----- 228
Db 312 EDYTLRKLLEPDELLIDRSTRDEEDQMIIRVDFTMIRRESGIFTGLVCVLHDVTEQEK 371
Qy 229 SFNQORQVEDASHELTPLIQIIOHNLNLIQRMGKDPAPVLESLSNISIEEMNRIKLVE 288
Db 372 NERERREFSVNSHELRTPLTSMRSYIEALSEGAWENPEIAPNPLKVTLEETDMIRMIN 431
Qy 289 ELLELTGQVNDISSEAQTVHINDEIRSRHSLKQL-----HPDYQFTDITSKNLEIKMK 344
Db 432 DLNLSRMSDGNTHLQLEYVNFNLELINFVLDRFDMTIENEQNKYKICREFTKRDLWYELD 491
Qy 345 PHOPEQLFLIFIDNAIKYDVKNKKIKVTKRLKNKQIIEITDHGIGIPEDQDFIPDRFY 404
Db 492 TDKVIQVLDNLNNAIKYSPDGGVITCLVETHNNVVFISDQGLGIPKDKLGKVERFY 551
Qy 405 RVDKSRSSQGGGLSGISIAQKIQLNGSGSIKSEINKGTTFKI 449
Db 552 RVDKARARAQGGTGLGLAISKEVIRAHNGSIWVESTEGSGTFFVI 596

RESULT 12
US-10-282-122A-57929
; Sequence 57929, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zvekind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: EUIRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 57929
; LENGTH: 611
; TYPE: PRT
; ORGANISM: Enterococcus faecium
US-10-282-122A-57929

Query Match
Best Local Similarity 17.8%; Score 408; DB 12; Length 611;
Matches 131; Conservative 114; Mismatches 184; Indels 186; Gaps 13;

Qy 3 KRKLK-----NNMIIVTMTFTVIFLFLCIIII-----FFLKD----- 35
Db 2 KRKVRFRSVNFKIA---ITFILILISIEIICAYFIRGLERSTINTFKDNMNTVESLA 58
Qy 36 -----TLHNSLDDAERSSSDINNLPHSKPVKDISALDNLNLSGNFQEIILIIYDEHNKLF 90
Db 59 TTISPENLRKDNADDEVANIKRFIENSATSDII-----EIRVDEKG--II 104
Qy 91 ETSNDNTVVERGEYEHRYFD--RVIKKRYKGIE-----YLIKEPI--TTQDFKGYSLLI 141
Db 105 RGTTDVNEQSAVGKKNYVDINDFTTKRYVALDNDKRVNINVPILSPGDAVIGALYV 164
Qy 142 HS-LENYDNIVKSLYIILAFGVITATISYVSTOITKPLVLSNKNMIEIRDGFQ 200
Db 165 KSNIEQKYEINNTAVIFFTASLIAAFISMIVSVLVARSITOPIGEMREQAIRIARGDYS 224
Qy 201 NKLQNLNTNYEEDNLANTFNMMSQIEESFN----- 231
Db 225 RKVKVH-GQDELQGLADTFNQLAERIEEAQDTWEARNRLDSVLSHMTDGVATDRRGKV 283
Qy 232 ----- 231
Db 284 ITINDWAVSLDDIKNEEAIGQSILTLDDIEEYTLRKLEDDPNMVLDRSTSTLESQMI 343
Qy 232 -----QQRQFVEDASHELTPLIQIIOHNLNLI 258
Db 344 LRVDFSMIRRESGFTISGLVAVLHDVTEQKTERERREFSVNSHELRTPLTSMRSYIEAL 403
Qy 259 QRWGKDPAPVLESLSNISIEEMNRIILKVELELTKGVDNDISSEAQTVHINDEIRSR 318
Db 404 SEGAWKDEEVAENFLKVTLEETDMIRMINDLNLSRMDTGNFQLEYVNFNLELVNVL 463
Qy 319 HSLKQL-----HPDYQFTDITSKNLEIKMKPHQFQELFLIFIDNAIKYDVKNKKIKVTKR 374
Db 464 DRFDMVNGQEKYKIRREFTORLWVEVDTKLIQVVDNINMNAIKYSPDGGTITCRLL 523
Qy 375 LKNQKQIIBITHGIGIPEDQDFIPDRFYRVKSRSSQGGGLSGISIAQKIQLNGGS 434
Db 524 ETHNNVVLSTIDQGLGIPKDKLNRVFERFYVDKARARAQGGTGLGLAISREVIAHRA 583
Qy 435 IKIKSEINKGTTFKI 449
```


Qy	203	LQUNTYEIDNLANTFNMSQIBEE	-----	228
		::: :::		
Db	229	VKIYGN-DEIGELALAFNNLSKRVQEAQANESKRRLLDSVITHMSDGIATDRGRIRI	287	

Qy	229	-----SF-----	230	
Db	288	VNDMALMKLGMAKEDIIGYMLSVLSDEPKLEBEOENNDSPFLDLINEEGLIARVNES	347	

Qy	231	-----NOOROFVEDASHELRLPLQIITQGHNLIIQRWKK	264	
		::: ::: :::		
Db	348	TIVQETGFTGYIAVLHDVTBQQQVERERREFVANVSHELRLPLTSNNSVIALEESGANK	407	

Qy	265	DPAVUEESUNISIEENRIIKVLBELLTGKD VNDIJSSEAQTVHINDEIRSIHSLKOL	324	
		::: ::: ::: :::		
Db	408	DEELAPQFLSVTRTEERMIRLVNDLLQLSMDNESDQINKEIIDFNMFINKIINRHEMS	467	

Qy	325	HPDYQFDTLTSKNLEIKMKHQPEQLFLIFIDNAIKVDKNKKIK--VKTRLKNQKII	382	
		::: ::: ::: ::: :::		
Db	468	AKDITFIROIPKKTIFTBFDPKMTQVFDNVITAMKYSRGDKRVEPHVKONPLYNRMTI	527	

Qy	383	EITDHGIGIPEEDQDPIFDRFVRVDKRSRSGQGNGLSLAOKIIOLNGGSIKIKSEIN	442	
Db	528	RIKDNGIGIPINKVDKIFDRFVRVDKARTKMGGTGLGLAISKEIVEAHNGRIWANSVEG	587	

Qy	443	KGTFKI	449	
Db	588	OGTSIFI	594	

RESULT 15

```

US-10-084-205-70
; Sequence 70, Application US/10084205
; Publication No. US20030049648A1
; GENERAL INFORMATION:
; APPLICANT: Choi, Gil
; TITLE OF INVENTION: 37 Staphylococcus aureus Genes and Polypeptides
; FILE REFERENCE: PB515P1
; CURRENT APPLICATION NUMBER: US/10/084,205
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: PCT/US00/23773
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: 60/151,933
; PRIOR FILING DATE: 1993-09-01
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 70
; LENGTH: 608
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-10-084-205-70

```

Query Match	17.4%	Score	397.5	DB	14	Length	608
Best Local Similarity	22.1%	Pred. No.	1.9e-22				
Matches	134	Conservative	105	Mismatches	187	Indels	181
Gaps	14						
Qy	4	RKLNNWIIIVTMTITFTVITFLCIIIFELKDTLHNSLDDAER-----	47				
Db	8	QSLHTKFIIVVVL-----IIIGMQLIGLYFTNNLEKLLDNFKNIIYQAKQLEISIEKV	63				
Qy	48	-----SSSDINNLF-----HSKPVKDISALDNLASLGNFQEIIVYDEHNK--LFETS	93				
Db	64	YDEKGSVNAOKDIOQLLSSEYANROEIGEIRPIDKD-----QIIITATTKOSRSLINQKA	117				
Qy	94	NDNTVR--VEPGYBHRFYDVRVKKRYKG-----IEVLIIEKEPITTDQFKGYSLLH	142				
Db	118	NDSSVQKALSQSQN-----DHLILKDYGGKDRVWVYNI PVKVDKVICNIYIES--KIN	171				
Qy	143	SLENYDNIVKSLYIIALAFGVIAITIIATTSYVFESTQITKPLVSLSNMKMIEIRRDGFQNK	202				
Db	172	DVYNQLNNINQIFVGTG---ISLLITVILGFFIARTTIPITDMRNQTVEMSRGNRTOR	228				
Qy	203	LQNTNVEEIDNLTANTENNMWSQEE-----	228				

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 4, 2004, 18:03:18 ; Search time 32.9836 Seconds
(without alignments)
705.906 Million cell updates/sec

Title: US-09-006-627-2
Perfect score: 2391
Sequence: 1 MTRKRLRNWIIVTMTIFV.....GGSIKIKSEINKGTFPKIIF 451

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents_AA*
1: /cgn2_6/ptodata/2/iaa/5A.COMB.pdp:*
2: /cgn2_6/ptodata/2/iaa/5B.COMB.pdp:*
3: /cgn2_6/ptodata/2/iaa/6A.COMB.pdp:*
4: /cgn2_6/ptodata/2/iaa/6B.COMB.pdp:*
5: /cgn2_6/ptodata/2/iaa/PTUS.COMB.pdp:*
6: /cgn2_6/ptodata/2/iaa/backfiles.pdp:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2291	100.0	451	US-09-642-000-2	Sequence 2, Appli
2	1640.5	71.6	471	US-09-134-001C-3150	Sequence 3150, Ap
3	625	27.3	521	US-09-107-532A-6431	Sequence 6431, Ap
4	594.5	25.9	375	US-09-134-000C-3962	Sequence 3962, Ap
5	422.5	18.4	583	US-09-272-414-2	Sequence 2, Appli
6	409.5	17.9	628	US-09-107-532A-7123	Sequence 7123, Ap
7	401	17.5	591	US-09-134-001C-3228	Sequence 3228, Ap
8	400.5	17.5	618	US-09-134-000C-4522	Sequence 4522, Ap
9	395	17.2	608	US-09-310-275-2	Sequence 2, Appli
10	395	17.2	608	US-09-082-077-5	Sequence 5, Appli
11	391.5	17.1	567	US-09-134-001C-3762	Sequence 3762, Ap
12	384	16.8	487	US-09-107-532A-5731	Sequence 5731, Ap
13	382	16.7	497	US-09-134-000C-4486	Sequence 4486, Ap
14	374.5	16.3	614	US-09-134-001C-3061	Sequence 3061, Ap
15	365.5	16.0	477	US-09-107-532A-5448	Sequence 5448, Ap
16	359	15.7	422	US-09-634-238-226	Sequence 226, App
17	351.5	15.3	464	US-09-543-681A-8192	Sequence 8192, Ap
18	348.5	15.2	352	US-09-107-532A-4211	Sequence 4211, Ap
19	338.5	14.8	559	US-09-328-352-4543	Sequence 4543, Ap
20	336	14.7	422	US-09-634-238-227	Sequence 227, App
21	330.5	14.4	918	US-08-843-530B-35	Sequence 35, Appl
22	319	13.9	460	US-09-134-001C-4107	Sequence 4107, Ap
23	317.5	13.9	480	US-09-489-039A-8863	Sequence 8863, Ap
24	317.5	13.9	491	US-09-489-039A-10199	Sequence 10199, A
25	315.5	13.8	585	US-09-107-532A-4126	Sequence 4126, Ap
26	311.5	13.6	350	US-09-094-103-6	Sequence 6, Appli
27	311.5	13.6	350	US-09-080-963-2	Sequence 2, Appli

Sequence 2, Appli
Sequence 27090, A
Sequence 29359, A
Sequence 68, Appl
Sequence 68, Appl
Sequence 25136, A
Sequence 8190, Ap
Sequence 7963, Ap
Sequence 5663, Ap
Sequence 8324, Ap
Sequence 4543, Ap
Sequence 3692, Ap
Sequence 2158, A
Sequence 6055, Ap
Sequence 5637, Ap
Sequence 31650, A
Sequence 6261, Ap
Sequence 24465, A

ALIGNMENTS

RESULT 1

US-09-642-000-2
; Sequence 2, Application US/09642000
; Patent No. 6548281
; GENERAL INFORMATION:
; APPLICANT: Wallis, Nicola G.
; Shilling, Lisa K.
; Mooney, Jeffrey L.
; Debouck, Christine
; Zhong, Yiyi
; Jaworski, Deborah D.
; Wang, Min
; Throup, John P.
; TITLE OF INVENTION: Histidine Kinase
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Dechert, Price & Rhoads
; STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103-2793
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/642,000
; FILING DATE: 18-Aug-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/006,627
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Falk, Stephen T
; REGISTRATION NUMBER: 36,795
; REFERENCE/DOCKET NUMBER: GM10127
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-994-2488
; TELEFAX: 215-994-2222
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 451 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-642-000-2

Query Match 100.0%; Score 2291; DB 4; Length 451;
Best Local Similarity 100.0%; Pred. No. 4.1e-185;
Matches 451; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MTKRKLNNWIIIVTTMTFTVTFIPLFCLIIIFFLKDTLHNSLDDAERSSSDINNLFHSPK 60
Db 1 MTKRKLNNWIIIVTTMTFTVTFIPLFCLIIIFFLKDTLHNSLDDAERSSSDINNLFHSPK 60
Qy 61 VKDISALDLNASLGNFQEIYDEHNHKKLPFETSDNTRVPEPGYEHRYFDRVVKRYKGI 120
Db 61 VKDISALDLNASLGNFQEIYDEHNHKKLPFETSDNTRVPEPGYEHRYFDRVVKRYKGI 120
Qy 121 EYLIKEPITTDQPKGYSLIHLNSLNDVNTVKSIIIALAFGIATIIITISYVFSTQI 180
Db 121 EYLIKEPITTDQPKGYSLIHLNSLNDVNTVKSIIIALAFGIATIIITISYVFSTQI 180
Qy 181 TKPLVSLNKMIEIRRDGFQNKQLNTNYEIDNLANTFNEMMSQIESFNQORQFVEDA 240
Db 181 TKPLVSLNKMIEIRRDGFQNKQLNTNYEIDNLANTFNEMMSQIESFNQORQFVEDA 240
Qy 241 SHELRTPLQIIQGHNLNLIQWKKDPKPAVLESLSNIEENNRILKLVLELLELTGKDVND 300
Db 241 SHELRTPLQIIQGHNLNLIQWKKDPKPAVLESLSNIEENNRILKLVLELLELTGKDVND 300
Qy 301 ISSEAQTVDHDEIRSRHSLKQLHPDYQFDTLTSKNLEIKMKPHQFEQLFLIFIDNAI 360
Db 301 ISSEAQTVDHDEIRSRHSLKQLHPDYQFDTLTSKNLEIKMKPHQFEQLFLIFIDNAI 360
Qy 361 KYDVNKKIKVTRKLNKKQKIETDGHGIGIPEDQDFIDRFYRVDKSRSRQGGNGLG 420
Db 361 KYDVNKKIKVTRKLNKKQKIETDGHGIGIPEDQDFIDRFYRVDKSRSRQGGNGLG 420
Qy 421 LSTAQKIQLNGSGSIKSEINKGTTFKIIIF 451
Db 421 LSTAQKIQLNGSGSIKSEINKGTTFKIIIF 451

RESULT 2
US-09-134-001C-3150
; Sequence 3150, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 3150
; LENGTH: 471
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3150

Query Match 71.68; Score 1640.5; DB 4; Length 471;
Best Local Similarity 69.9%; Pred. No. 3.2e-130;
Matches 316; Conservative 68; Mismatches 67; Indels 1; Gaps 1;

Qy 1 MTKR-CLRNNWIIIVTTMTFTVTFIPLFCLIIIFFLKDTLHNSLDDAERSSSDINNLFHSPK 59
Db 16 MIKQKLYKXWMLTITLTITLILFCLIIIFFLKDTLHNSLDDAERSSSDINNLFHSPK 75
Qy 60 PVKDISALDLNASLGNFQEIYDEHNHKKLPFETSDNTRVPEPGYEHRYFDRVVKRYKGI 119
Db 76 SLSDISALDLNASLGNFQEIYDKGRKLIQTSNDNTLAYDNKIDPKPHERIHQSRHG 135
Qy 120 IEYLIKEPITTDQPKGYSLIHLNSLNDVNTVKSIIIALAFGIATIIITISYVFSTQ 179

Db 136 INVLVTEPIRSKDFSGYSVLVHSLQYDNLVKSIIYVALAFLGIATIIITAGVSIFSSQ 195
Qy 180 ITKPLVSLNKMIEIRRDGFQNKQLNTNYEIDNLANTFNEMMSQIESFNQORQFVED 239
Db 196 ITKPIVTSMKNQIIRRDGFQNKLELTNTTYEETDNLIDTFNEMMYQIESFNQORQFVED 255
Qy 240 ASHELRTPLQIIQGHNLNLIQWKKDPKPAVLESLSNIEENNRILKLVLELLELTGKDVN 299
Db 256 ASHELRTPLQIIQGHNLNLIQWKKDPKPAVLESLSNIEENNRITKLVLELLELTGKDVN 315
Qy 300 DISSEAQTVDHDEIRSRHSLKQLHPDYQFDTLTSKNLEIKMKPHQFEQLFLIFIDNA 359
Db 316 HNVLECEVDVNSEIQSRVKSLOHLHPDYTFETHLATKPIQLKINRHOEQLLIIFIDNA 375
Qy 360 IKYDVNKKIKVTRKLNKKQKIETDGHGIGIPEDQDFIDRFYRVDKSRSRQGGNGL 419
Db 376 MKYDTEHKHKIKVTQLKNKWMIDITDHGWIPIKADLEFIDRFYRVDKSRSRQGGNGL 435
Qy 420 GLSIAQKIQLNGSGSIKSEINKGTTFKIIIF 451
Db 436 GLSIAEKIVQLNGSGMIQVSESLQKYTFKISF 467

RESULT 3
US-09-107-532A-6431
; Sequence 6431, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD/ROM ISO9660
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 6431:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 521 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...521
; SEQUENCE DESCRIPTION: SEQ ID NO: 6431:

US-09-107-532A-6431

Query Match 27.3%; Score 625; DB 4; Length 521;
Best Local Similarity 31.7%; Pred. No. 1.7e-44;
Matches 158; Conservative 108; Mismatches 165; Indels 68; Gaps 14;

Qy 3 KRKRN-----NWIIVTMTFTVIFLPCIIIFLKDTHLSELDAAERSSDI-----52
Db 29 KRELKGPLATIKWAFASFFIVFTTIPAVITYKSSINLIIVAKERNVVERTISEVASRLA 88
Qy 53 -----NNLFHSPVKVKSALDNLASL--GNFQEI-----IYDE 84
Db 89 NSDEBLTVNRYNL--TKANDGDEVVDHSLVEGNLMDKIDSFIAELGQPSMDLYVDL 146
Qy 85 HNNKLFETSNDN-----TVRVEPGVEHRYFDRVI-----KKRYGIEVLIIEKEPITTTQDFK 135
Db 147 DENLIFKTHENSRLIQTTRKVP-----TIVLEGKTGFLSIQPIYSK--GTREKI 195
Qy 136 GYSLIHSLHNDYDNIVKSLYIIAALAFGVIATIIATISYVFTSTQITKPLVSLSNKMEIR 195
Db 196 GYAQSVELSFYDIRNKLTLTLVLEVVSILSLGFLFLSYFLKPLKVLRTDMTIR 255
Qy 196 RDGFQNKI--QLANTVEEIONLANTENMSQIEESFNQORQFVEDASHELRTPLQIIQ 252
Db 256 KDP-QSDIHMEIDTN-DELADLAEIFNEMLDRLMYIEQOEQFVEDVSHELTPVALIE 313
Qy 253 GHLNLIQKWKGPAPVLEESLNIEEMNRILKVEELLELTGKDVNDISSEAQTVHND 312
Db 314 GHLNLIQKWKGPAPVLEESLNIEEMNRILKVEELLELTGKDVNDISSEAQTVHND 312
Qy 313 EIRSRHSLKQLHPDY--QFTDTLTSKMLETKMKPHQEPQLFIPIDNAIKYDVNKKIK 370
Db 374 VTYQVNNFKILYPDYVITLDDLT-KEVTLQIYRNHPEQLIIILDNNAVYSTRKEVH 432
Qy 371 VKTRLKKNKQKIIETDHGIGIPEDODFIDRFYRVKRSRSGGNGLGSLIAQKLIOL 430
Db 433 ISISSTLNEFFIAIQDFEGEIPKEDLNKIFNRFYRVKARARTKGGNGLGSLIAQKLVEN 492
Qy 431 NGGSIKIKSEINKGTTFKI 449
Db 493 YKGRILAESVVGQTFIRI 511

RESULT 4

US-09-134-000C-3962
; Sequence 3962, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3962
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-134-000C-3962

Query Match 25.9%; Score 594.5; DB 4; Length 375;
Best Local Similarity 35.8%; Pred. No. 4e-42;
Matches 129; Conservative 88; Mismatches 124; Indels 19; Gaps 7;

Qy 95 DNTVRVPGVEHRYFDRVIRKRYGIEVLIIEKEPITTTQDFKGYSLIHSLHNDYDNIVKSL 154
Db 16 DNYLLFEP-----FLIKRVFISVEPIFSKE---TREKIGYIQAQVELSSFVEIRNHL 64
Qy 155 YIIALAFGVIIATIIATISYVFTSTQITKPLVSLSNKMEIRRDGFQNKI---QLNTNVEE 211

Db 65 LLTLVLEVISLVSSVILGFILSSYFLKPLKVLRTDMDTIRKDP-QSDVHMEINTR-DE 122
Qy 212 IDNLANTFNEMMSQIEESFNQORQFVEDASHELRTPLQIIQGHNLNLIQKWKGPAPVLEB 271
Db 123 LADISEIFNEMLDNRMYIEQOEQFVEDVSHELRTPVAIMEGHLNLIQKWKGPAPVLEB 182
Qy 272 SLNISIEEMNRILKVEELLELTGKDVNDISSEAQTVHINDEIRSRHSLKQLHPDYQ-- 329
Db 183 SLKASLQEI SRMSLVQEMLDLSRAEQVDQYANERTDAKQVYVQVFNFLVYVPEFHT 242
Qy 330 FDTDLTSKNLEIKMKPHQEPQLFIPIDNAIKYDVNKKIKYKTRLNKQKIIETDHI 389
Db 243 LDDDLPTLPE-VELKIYRNHPEQLIIILDNNAIKYSTDRKEVHISIRTNWSEFIAVQDFGE 301
Qy 390 GIPEEDODFIDRFYRVKRSRSGGNGLGSLIAQKIIQLNGGSIKIKSEINKGTTFKI 449
Db 302 GITTEDELEKIFDRFYRVKARANKGGNGLGSLIAQKOLVENYKGRIDAESVLHQGTIFRI 361

RESULT 5
US-09-272-414-2
; Sequence 2, Application US/09272414A
; Patent No. 6238885
; GENERAL INFORMATION:
; APPLICANT: Wallis, Nicola G.
; TITLE OF INVENTION: Histidine Kinase
; FILE REFERENCE: GM10202
; CURRENT APPLICATION NUMBER: US/09/272,414A
; CURRENT FILING DATE: 1999-03-19
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 583
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-272-414-2

Query Match 18.4%; Score 422.5; DB 3; Length 583;
Best Local Similarity 23.7%; Pred. No. 2.4e-27;
Matches 137; Conservative 90; Mismatches 195; Indels 155; Gaps 12;

Qy 11 IIVTMTFTVIFLPCIIIFLKDTHLSELDAAERSSDINNLFHSK--PVKDISALD 68
Db 18 LIVTVTLILLISALITPMQYFTQET-ENAIREDARRISSIVEQ-SHNKSEAIKYSOTLI 75
Qy 69 LN-----ASLGNFQEIILYDEHNKLPFETSDNTVTRVPGVEHRYFDRVI 113
Db 76 ENPGGLMIINKHQRQSTASLSNKKQMLNEVNNHDFDVFDDKGSVTRNV-----TI 128
Qy 114 KRYKGYEYLIIKEPITTTQ--DFKGSY--LLIHSLENDYDNIVKSLYIIAALAFGVIIATIT 169
Db 129 KEGSSQTYILLGYPTKQKNSHSGVGYIKDKLSIEDTNNAITITITITVIFLIT 188
Qy 170 ATISYVFTSTQITKPLVSLSNKMEIRRDGFQNKQLQNTNVEEIDNLANTFNEMMSQIEB- 228
Db 189 TVFAFLSSRITKPLRLRQATRVSEGDYSYKPSVTTK-DEIGQLQAQNMSTEIEEH 247
Qy 229 -----SFOQORQ----- 235
Db 248 VDALSTSKNIRDSLINSWVEGLGINESRQIILSNKMANDIMDNIDEDAKAFLRLQIEDT 307
Qy 236 -----FVEDA 240
Db 308 FKSQKTEMERDLENARFVVTTSYIDKIEGGSGVGVVTVRDMTNEHNLDMKKDFANV 367
Qy 241 SHELRTPLQIIQGHNLNLIQKWKGPAPVLEESLNIEEMNRILKVEELLELTGKDVND 300
Db 368 SHELRTPLSLQGVTEISIVDGPTEPEIKESLAVLDESKRLNRLNELLNVARMDAEG 427
Qy 301 ISSEAQTVHINDEIRSRHSLKQLHPDYQFDTDLTSLQLEI-----KKKPHQFQLF 352
Db 428 LS-----VNKEVQPIAALLDKMKIKYRQQAADGLNMTFNYCKRWVSYDMRMDQVL 480

Db 128 Q--NITIKNGNSQSYLLGVPMKAQ--KGAQSHYSGVFIYKOLKSIEDTNNATIIILI 183
Qy 161 FGVIATITATISYVFSTQITKPLVSLNKMIEIRRDGFONKQLNTVVEIDNLANTFN 220
Db 184 TALLIFTIATIFAPFLSNRIKPLQLKTQAKV--SEGDYSGISVATKDEIGDLSRAPN 242
Qy 221 EMMSQIEE-----SFNQQRQ----- 235
Db 243 NNVBEIOEHKAISSSKNIRDTLNLMVGLGINNOREIILSNKMWADDIMRHIDDFSKE 302
Qy 236 ----- 235
Db 303 SIEQIEATFESQNEYLEINTRYVFISSYIDRIOTNRSRGVIMVDMTNEHNDQ 362
Qy 236 ----FVEDASHELRTPLQIIQGHNLIIQWGGKDPVAVLEESINISIEEMRIIKLVEELL 291
Db 363 MKKDPANVSHELRTPIQLQGYTESIVDGIPTEDIRSDLAIVLDESRLNLVNEELL 422
Qy 292 ELTKGDVNDISSEAQ---TVHINDEIRSIH--SLKQLHPDYQFDTLTSKNLEIKMKPKHQ 347
Db 423 NVARMDAEGLSVEKELQPIQHLDRKESKYMSEGLGLTWTDFSNNDQELWNYDM--DR 480
Qy 348 FEQLFLIFIDNAIKYDVNKKIKVTKLNKQKIIETDGHGIGIPEEDQFIEDRFYRVD 407
Db 481 MDQVLNLIDNATRYTQAGDSIKISIDEDSPNIIITDTGTGIAPEHLKQVDFRYKVD 540
Qy 408 KSRSRSGGNGLGLSIAQKIIQLNGSGSIKISEINKGTTFKI 449
Db 541 AARKEGKGTGLGLFICKRWIIEHGRIDVSEGLKGTSFII 582

RESULT 8

US-09-134-000C-4522
; Sequence 4522, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4522
; LENGTH: 618
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (8)-(8)
; OTHER INFORMATION: Amino acid 8 is Xaa wherein Xaa = any amino acid.
US-09-134-000C-4522

Query Match 17.5%; Score 400.5; DB 4; Length 618;
Best Local Similarity 20.7%; Pred. No. 1.9e-25;
Matches 123; Conservative 124; Mismatches 173; Indels 173; Gaps 15;
Qy 17 ITFTVIFLFLCIIII--PFLKDTLHNSLDDAERS-----SSDINNLFHSPVK 62
Db LSFILLLLIAIIGGVVIRE-LEATTISDFKNMDSQVQLSDMLSTQMSN-----K 77
Qy 63 DISALDNLASL-----GNFOEIIYDHNNKLPETSDNTVRVPGVEHYRFD--R 111
Db DLVRSDDANLKKALSDFSNADISEARIVDDKG--IIRATNDLQNNIIGKNDYRLND 135
Qy 112 VIKRYKIGIE-----YLIIEKPI---TTQDPKGYSLHLHLEN--YDNIVKSLYIIIALAFG 162
Db 136 FTSKKYQALNDKRVYVNV--QPIOSPTGETVIGVLVYKSNLENKIQEITNTASIFFTA-S 193
Qy 163 VIATITATISYVFSTQITKPLVSLNKMIEIRRDGFONKQLNTVVEIDNLANTFNEM 222

Db 194 IIAAAISIIIVTLIARSITKPIGEMREQAIRARGDYAGKVEVH--GKDELQGLAETFNQL 252
Qy 223 MSQIEE----- 228
Db 253 SERIEEAQETWEABERNRLDSVLTHMTDGVATDRRKVITINEMALSLLANVKNENVIOTS 312
Qy 229 ----- 228
Db 313 LLELLDIEDYTLAKLLEEPDELLIDRSTSDREEDQMIIRVDFTMIRRESGFIITGLVCVL 372
Qy 229 -----SFNQQRQFVEDASHELRTPLQIIQGHNLIIQWGGKDPVAVLEESINISIEEM 280
Db 373 HDVTEQEKNEREREFVSNVSHELRTPLTSRYSIEALSEGAWENPETAPFLKVTLEET 432
Qy 281 NRIIKLVEELLETKGDVNDISSEAQTVHINDEIRSIHSLKQL---HPDYQDPDTLTS 336
Db 433 DRMRINDLNLNRMDSGNTHLQLEVYVNFELINFLVDRFDMMIENEQKNYKICREFTK 492
Qy 337 KNLEIKMKPHQFEQLFLIFIDNAIKYDVNKKIKVTKLNKQKIIETDGHGIGIPEEDQ 396
Db 493 RDLAVELDTDKVIQVLDNLLNNAIKYSPDGGVITCRVETHNNVVFISDQGLGIPKKDL 552
Qy 397 DFIDRFYRVDKSRSRSGGNGLGLSIAQKIIQLNGSGSIKISEINKGTTFKI 449
Db 553 GKVFERYRVDKARARAGGTGLGLAISKEVIRAHNGSIWVESTEGETGTFYI 605
RESULT 9
US-09-310-275-2
; Sequence 2, Application US/09310275
; Patent No. 6194174
; GENERAL INFORMATION:
; APPLICANT: Throup, John P.
; APPLICANT: Palmer, Leslie M.
; APPLICANT: Van Horn, Stephanie
; APPLICANT: Warren, Richard L.
; APPLICANT: Burnham, Martin K. R.
; TITLE OF INVENTION: 636HK
; FILE REFERENCE: GM10222
; CURRENT APPLICATION NUMBER: US/09/310,275
; CURRENT FILING DATE: 1999-05-12
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 608
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-310-275-2
Query Match 17.2%; Score 395; DB 3; Length 608;
Best Local Similarity 22.2%; Pred. No. 5.4e-25;
Matches 133; Conservative 102; Mismatches 184; Indels 180; Gaps 14;
Qy 15 TMITFTVIFLFLC---IIFFLKDTLHNSLDDAER----- 47
Db 12 TKLVIVVVLIIIGMQLIIGLYFTNNLEKLLDNPKNITQYAKQLEISIEKVDKGSVN 71
Qy 48 SSDINNLF---HSKPVKDISALDNLASIGNFQEIIEIYDEHNNK--LFETSNDNTVR-- 99
Db 72 AQKDIQNLSEYANRQGEIRFIDKD-----QIIATTKQSNRSLNQKANDSSVQKA 125
Qy 100 VEPGYEHRYDRVITKKYKG-----LEYLIKEPITTDPKGYSLHLHLENVDNI 150
Db 126 LSLGQSN---DHLIKDYGKKORVWVYVNPVKVKVIGNIYES---KINDVYNQLNN 179
Qy 151 VKSLYIITALAFGIATITATISYVFSTQITKPLVSLNKMIEIRRDGFONKQLNTVVE 210
Db 180 INQIFIVGTA---ISLITVILGFFIARTITKPIIDMRNQTVMSRGNVYQRVKIYGN-D 235
Qy 211 EIONLANTFNMSQIEE----- 228
Db 236 EIGELALAFNLSKRVQEAQANTESEKRLDSVITHMSDGLIATDRGRIRIVNDMALKM 295

QY 229 -----SF----- 230
 Db 296 LGMAKEDIIGYMLSVLSLEDEFLKEIEIQENNDSEFLDLNEBGLIARVNFSTIVQETGF 355
 QY 231 -----NQORQFVEDASHELRTPLQIIQGHNLNLIQRMGKDPVAVLEES 272
 Db 356 VTGYIAVLHVDVTEQQOVEREREFVANVSHELRTPLTSMNSYIEALEEGAWKDEELAPQF 415
 QY 273 LNISIEEMNRIIKVLBEELTETKGVNDISSEBAQTWHINDEIRSRHSLKQLHPDYQFDT 332
 Db 416 LSVTREETERMIRLVNDLLQSLKMDNESQINKEIIDFNFNKINRHEMSAKDTTFIR 475
 QY 333 DLTSKNLEIKMKPHQEPQLFIIDNAIKYDVKNKKIK--VKTRLNKKQKIIEITHGIG 390
 Db 476 DIPKKTIFTEPDQKMTQVDFNVITNAMYSGDKRVEFHVKQNPVPLNRTWIRIKONGIG 535
 QY 391 IPBEDODFIIDRFYRDKSRSSRSGGNGLSIAQKIIQIOLNGSGSIKIKSINKGTTFKI 449
 Db 536 IPINKVDKIFDRFYRDKARTKMGGTGLGLAISKEIVEAHNGRIWANSVEGQGTSTIFI 594

RESULT 10

US-09-082-077-5
 ; Sequence 5, Application US/09082077
 ; Patent No. 6514746
 ; GENERAL INFORMATION:
 ; APPLICANT: Microcide Pharmaceuticals, Inc.
 ; TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS HISTIDINE PROTEIN KINASE
 ; FILE REFERENCE: SR Silverstein Microcide 234/067US
 ; CURRENT APPLICATION NUMBER: US/09/082,077
 ; CURRENT FILING DATE: 1998-05-20
 ; PRIOR APPLICATION NUMBER: US 60/009,102
 ; PRIOR FILING DATE: 1995-12-22
 ; PRIOR APPLICATION NUMBER: PCT/US97/23912
 ; PRIOR FILING DATE: 1997-12-23
 ; PRIOR APPLICATION NUMBER: US 08/713,718
 ; PRIOR FILING DATE: 1996-09-13
 ; NUMBER OF SEQ ID NOS: 5
 ; SOFTWARE: Patent in Ver. 2.1
 ; SEQ ID NO 5
 ; TYPE: PRT
 ; LENGTH: 608
 ; ORGANISM: Staphylococcus aureus
 US-09-082-077-5

Query Match 17.2%; Score 395; DB 4; Length 608;
 Best Local Similarity 22.2%; Pred. No. 5.4e-25;
 Matches 133; Conservative 102; Mismatches 184; Indels 180; Gaps 14;
 QY 15 TMTFTVTFIFCL---IIIFLKDTHLSBELDAAER----- 47
 Db 12 TKLVVVVLLIIIGMQLIIGLYFTNNLEKELDNFKKNTQYAKOLEISIEKVYDEKGSVN 71
 QY 48 SSSDINNLF---HSRPVKDISALDNASLGNQEIILYDEHNNK--LPFETSDNTRV-- 99
 Db 72 AQKDQIQLLSEYANROBGEIRFIDKD-----QIIATTKQNSRLNQKANDSSVQKA 125
 QY 100 VEPGEYHRYFDRVVKRYKG-----IEYLIKEPITQDPFGYSLLHLSLENDNI 150
 Db 126 LSLGQSN---DHLLKDYGGKDRVWVYNTPVKVDKKVIGNIYES----KINDVYQQLNN 179
 QY 151 VKSLYIIALAFGVIATITATISYVFSTQITKPLVSLNKMIEIRDRGFQNKQLANTNYE 210
 Db 180 INQIFIVGTA---ISLLITVILGFFARTITKPIDTMRNQTVMESRGNVYQRVKIYGN-D 235
 QY 211 EIDNLANTFNMMSQIEE-----IEAHENKIAQVITNLLSNALNYSSEDNKVIVRYRNDNKVYLEI 228
 Db 236 EIGELALAFNLSKRVQEAQANTESEKRLDSVITHMSDGIATDRGRIRIYNDMALKM 295
 QY 229 -----SF----- 230

Db 296 LGMAKEDIIGYMLSVLSLEDEFLKEIEIQENNDSEFLDLNEBGLIARVNFSTIVQETGF 355
 QY 231 -----NQORQFVEDASHELRTPLQIIQGHNLNLIQRMGKDPVAVLEES 272
 Db 356 VTGYIAVLHVDVTEQQOVEREREFVANVSHELRTPLTSMNSYIEALEEGAWKDEELAPQF 415
 QY 273 LNISIEEMNRIIKVLBEELTETKGVNDISSEBAQTWHINDEIRSRHSLKQLHPDYQFDT 332
 Db 416 LSVTREETERMIRLVNDLLQSLKMDNESQINKEIIDFNFNKINRHEMSAKDTTFIR 475
 QY 333 DLTSKNLEIKMKPHQEPQLFIIDNAIKYDVKNKKIK--VKTRLNKKQKIIEITHGIG 390
 Db 476 DIPKKTIFTEPDQKMTQVDFNVITNAMYSGDKRVEFHVKQNPVPLNRTWIRIKONGIG 535
 QY 391 IPBEDODFIIDRFYRDKSRSSRSGGNGLSIAQKIIQIOLNGSGSIKIKSINKGTTFKI 449
 Db 536 IPINKVDKIFDRFYRDKARTKMGGTGLGLAISKEIVEAHNGRIWANSVEGQGTSTIFI 594

RESULT 11

US-09-134-001C-3762
 ; Sequence 3762, Application US/09134001C
 ; Patent No. 6380370
 ; GENERAL INFORMATION:
 ; APPLICANT: Lynn Doucette-Stamm et al
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
 ; FILE REFERENCE: GTC-007
 ; CURRENT APPLICATION NUMBER: US/09/134,001C
 ; CURRENT FILING DATE: 1998-08-13
 ; PRIOR APPLICATION NUMBER: US 60/064,964
 ; PRIOR FILING DATE: 1997-11-08
 ; PRIOR APPLICATION NUMBER: US 60/055,779
 ; PRIOR FILING DATE: 1997-08-14
 ; NUMBER OF SEQ ID NOS: 5674
 ; SEQ ID NO 3762
 ; LENGTH: 567
 ; TYPE: PRT
 ; ORGANISM: Staphylococcus epidermidis
 US-09-134-001C-3762

Query Match 17.1%; Score 391.5; DB 4; Length 567;
 Best Local Similarity 27.2%; Pred. No. 9.7e-25;
 Matches 116; Conservative 84; Mismatches 136; Indels 91; Gaps 13;
 QY 43 DDAERSSDINNLFHSPVKDISALDNASLGNF-QEIIYDEH-----NNKLFETSND 95
 Db 211 NDLARRLQKLN-----SQIKSNRLKTTLENIPSSVLMDKHGEIVVANHAYYQVFN 264
 QY 96 NTVRVEPGYEHRYFDRVVKRYKG-----IEYLIKEPITQDPFGYSLLHLSLENDNI 150
 Db 265 DQM-----VENKSYIGFIDDSIEKLI-----ESFRTKVKYEQLEVAINN 305
 QY 151 VKSLYIIALAFGVIATITATISYVFSTQITKPLVSLNKMIEIRDRGFQNKQLANTNYE 210
 Db 306 VHTKYP-----DVSCI-----PILTKSKQLQ-----GWVVLHDTNLQ 340
 QY 211 EIDNLANTFNMMSQIEESFNQORQVEDASHELRTPLQIIQGHNLNLIQRMGKDPVAVLE 270
 Db 341 KLENL-----RREFVANVSHELKTPITSIKGFAETLIEGAKNDEQSLD 383
 QY 271 ESINISIEEMNRIIKVLBEELT-----KGDVNDISSEBAQTWHINDEIRSRHSLKQL 324
 Db 384 MELNITLESNRIBESLVTLDDLSHTEQOKELEINYNLSELAINIDNLQAYN-KRI 442
 QY 325 HPDQPDFTLTSKNLEIKMKPHQEQFLIFIDNAIKYDVKNKKIKVKTRLNKKQKIIEI 384
 Db 443 KQSEIEKQVI-----IEAHENKIAQVITNLLSNALNYSSEDNKVIVRYRNDNKVYLEI 497
 QY 385 THGIGIPEEDQDFPDRFYRDKSRSSRSGGNGIGLSIAQKIIQIOLNGSGSIKIKSINKG 444
 Db 498 QDYGIGISDQKRIPEFYRDKARSRDSGGTGLGLSITKTHIVEAHNGRIDVKSAPGK 557


```

QY 415 GGNGLGLSTAKLIQNLGGSGIKSEINKTKTTFKIIF 451
Db 453 GESGLGLPIVQVRLHKGKINVESELGKGTTFIISF 489

RESULT 14
US-09-134-001C-3061
; Sequence 3061, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134, 001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 3061
; LENGTH: 614
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3061

Query Match 16.3%; Score 374.5; DB 4; Length 614;
Best Local Similarity 26.3%; Pred. No. 2.9e-23;
Matches 124; Conservative 88; Mismatches 167; Indels 93; Gaps 14;

QY 10 WIIVTMTITVTFLFCLIIIFLKTLSNELDRAESSDINNLFHSPKVDISALDL 69
Db 190 FIVGTALSLFTV-----ILGFFIARTI-----TKPITDMRNQTV 224

QY 70 NASLGNP-QSIIIVYDEHNKLFETS---NDNTRVPEPGY-----EHRFDRVFKRYKGI 120
Db 225 EMSKGNVQTVKLYG--NDEIGELALAFNNLSKRVQEAQANTSEKRLDSVITHMSDGI 282

QY 121 -----EYLIIKEPITTPQPKGYSL--IHSLEN-----YDNIVKSLYIAL 159
Db 283 LATDRGRVRIANDMALMGLAKEDVIGYMLGVNLNENEPSELSIEQNSDSFLLDINE 342

QY 160 AFGVIATITATISYFSTQITKPLVSLSNKMEIIRDDGQNKQLNTNVEEDINLNTF 219
Db 343 EGGIARVNFTSI-----VOETGFV-----TGVIIV-----L 369

QY 220 NEMMSQIESFNQORQVEDASHELRTPLQIIQGHNLNIQRWKKDPVAVLEESLNTSIE 279
Db 370 HDVTEQ-QQVERERREFVANVSHLRTPLTSMNSYIEALEEGAWQKELAPSLSVTREE 428

QY 280 MNRRIKVLLELBTKGVDNDISSEAQTVDHINDEIRSRHSLKQLHPDYQFDTLTSKNL 339
Db 429 TERMIRLVNDLQLSKMNDSDQITKEIIDFNMFINKIINRHEMAAKDTTFVREIPQOTI 488

QY 340 EIKMKPHQFQLFLIFDINAIVKVDVKNKKIK--VKRTLKNKKQKLIETDGHGIGIPEDQD 397
Db 489 FAEIDPKMTQVDNVTINAMKYSRGEKRVFHVQKNALYNRTIRIKNGIGIGIPINKVD 548

QY 398 FIDRFYRVDKSRSSRCGGNGLGLSTAKLIQNLGGSGIKSEINKTKTTFKI 449
Db 549 KIFDRFVRVDKATRMGGTGLGLAISKEIVAEHNGRIWANSVEGGQTSIFI 600

RESULT 15
US-09-107-532A-5448
; Sequence 5448, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310

```

```

CORRESPONDENCE ADDRESS:
ADDRESSEE: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
City: Waltham
STATE: Massachusetts
COUNTRY: USA
Zip: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD/ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Ariniello, Pamela Denke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 5448:
SEQUENCE CHARACTERISTICS:
LENGTH: 477 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...477
SEQUENCE DESCRIPTION: SEQ ID NO: 5448:
US-09-107-532A-5448

Query Match 16.0%; Score 365.5; DB 4; Length 477;
Best Local Similarity 25.3%; Pred. No. 1.2e-22;
Matches 104; Conservative 83; Mismatches 145; Indels 79; Gaps 9;

QY 53 NNLFSKPV---KDISALDNASLGNFQBIIIYDEHNKLFETSNDNTRV-----VEPGY 104
Db 124 NDRFNTQVILWSKNGEILLKAAIGGRFNQI-----KELKMDTKNLNTIQLNELNEDSGS 177

QY 105 EHRYPDRVFKRYKGYEYLIIEKEPITTPQPKGYSLIHSLENVDNIVKSLYIALAFGVI 164
Db 178 SALSFRSITTKAE-----TNQDGVAYIQILENTNQIKNSLATFRTIMILCMIV 225

QY 165 ATITATISYFSTQITKPLVSLSNKMEIIRDDGQNKQLNTNVEEDINLANTFNEMMS 224
Db 226 FWLSIGISYLSLSMRFIIL----- 247

QY 225 QIEESFNQORQVEDASHELRTPLQIIQGHNLNIQRWKKDPVAVLEESLNI--SIEEMNR 282
Db 248 ----SWRRQKEFVENASHELRTPLTIQNSLEHL--FTKPNHTIIEESSIAQALSETRR 301

QY 283 IYKLVLELLELTGVDNDISSEAQTVDHINDEIRSRHSLKQL-----HPDYQFDTLTSKN 338
Db 302 LTGLTSDLTITARNDSQQLSKQMINTQSYEQLVKPFQEMAIIDKGEFILHNDAT--- 358

QY 339 LEIKMKPHQFQLFLIFDINAIVKVDVKNKKIKVTKRLKNKKIETDGHGIGIPEDQDF 398
Db 359 LEVPIDEKKIHQVLVILLNALKYTNPEDKIILDSSVNHKEWLIHVKNVTGSPSIEEDQAR 418

QY 399 IFDRFYRVDKSRSSRCGGNGLGLSTAKLIQNLGGSGIKSEINKTKTTFKI 449
Db 419 IFDRFYRVQDSRAKETGGYGLGLATKQVIEQHHGQISVSNVLPQGVDFKV 469

```

Search completed: October 4, 2004, 18:11:17
Job time : 36.9836 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 4, 2004, 17:52:23 ; Search time 118.472 Seconds
(without alignments)
1075.607 Million cell updates/sec

Title: US-09-006-627-2
Perfect score: 2291
Sequence: 1 MTRKRLRNWIIVTMTIFV.....GGSIKIKSEIKNGTTFKIIIF 451

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_29Jan04.*

- 1: Geneseqp1980s.*
- 2: Geneseqp1990s.*
- 3: Geneseqp2000s.*
- 4: Geneseqp2001s.*
- 5: Geneseqp2002s.*
- 6: Geneseqp2003as.*
- 7: Geneseqp2003bs.*
- 8: Geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2291	100.0	451	2	AAY28601 Histidine
2	2291	100.0	451	6	ABU42408 Protein e
3	2291	100.0	451	6	ABM73032 Staphyloc
4	1640.5	71.6	471	5	ABP38305 Staphyloc
5	1069	46.7	284	4	ABG82166 S. epider
6	1069	46.7	284	4	ABG82848 S. epider
7	703.5	30.7	483	5	ABB48292 Listeria
8	703.5	30.7	483	6	ABU33012 Protein e
9	625	27.3	521	7	ADC96804 E. faeciu
10	606.5	26.5	491	5	ABB54939 Lactococc
11	589.5	25.7	506	6	ABD09380 Alloiococ
12	552.5	24.1	330	6	ABD09378 Alloiococ
13	532.5	23.2	500	5	ABP26467 Streptoco
14	515	22.5	501	5	ABP27834 Streptoco
15	515	22.5	501	5	ABP29835 Streptoco
16	464.5	20.3	482	6	ABU24047 Protein e
17	452.5	19.8	441	5	ABU54320 Lactococc
18	426	18.6	585	6	ABU43406 Protein e
19	425.5	18.6	579	4	AAU33824 Staphyloc
20	425.5	18.6	583	4	AAU37004 Staphyloc
21	424.5	18.5	583	6	ABM71083 Staphyloc
22	424	18.5	613	6	ABU17734 Protein e
23	422.5	18.4	583	3	ABU23444 Staphyloc
24	422.5	18.4	588	6	ABU16260 Protein e
25	421.5	18.4	628	6	ADB07016 Alloiococ

26	419.5	18.3	603	6	ADB07014 Alloiococ
27	413.5	18.0	609	4	AAU35156 Enterococ
28	412.5	18.0	612	6	ABU14658 Protein e
29	409.5	17.9	628	7	ADC97496 E. faeciu
30	408	17.8	611	6	ABU30005 Protein e
31	401	17.5	589	6	ABU42686 Protein e
32	401	17.5	591	5	ABP38383 Staphyloc
33	400	17.5	591	5	ABP38383 Staphyloc
34	397.5	17.4	608	4	ABU00862 Listeria
35	396.5	17.3	449	3	AAU00862 S. aureus
36	396.5	17.3	449	3	AAU00862 S. aureus
37	396.5	17.3	449	6	ABU01655 S. pneumo
38	395.5	17.3	449	6	ABP81398 Streptoco
39	395.5	17.3	449	4	AAU37673 Streptoco
40	395.5	17.3	449	6	AAU37966 Streptoco
41	395	17.2	608	2	AAU06431 Staphyloc
42	395	17.2	608	4	AAU37406 Staphyloc
43	395	17.2	608	4	AAU34181 Staphyloc
44	395	17.2	608	4	AAU37076 Staphyloc
45	395	17.2	608	4	AAU37097 Staphyloc

ALIGNMENTS

RESULT 1
AAY28601
ID AAY28601 standard; protein; 451 AA.

AC AAY28601;

DT 01-OCT-1999 (first entry)

DE Histidine kinase protein from Staphylococcus aureus.

KW Histidine kinase; Two component signal transduction system; TCSTS;
virulence; bacterial pathogenesis; Lactococcus lactis cremoris; KlnA;
autophosphorylate; cognate response regulator; genetic immunisation;
antibacterial agent; screening; Helicobacter pylori; gastric ulcer.

OS Staphylococcus aureus.

PN WO9936508-A1.

PD 22-JUL-1999.

PF 12-JAN-1999; 99WO-US0000610.

PR 13-JAN-1998; 98US-00006627.

PA (SMIK) SMITHKLINE BEECHAM CORP.

PI Wallis NG, Shilling LK, Mooney JL, Debouck C, Zhong Y;

Jaworski DD, Wang M, Throup JP;

WPI; 1999-444390/37.

N-PSDB; AAX90601.

Novel histidine kinase polynucleotides and polypeptides used to screen for antibacterial compounds.

Claim 1; Page 6; 43pp; English.

The present sequence is a histidine kinase protein from Staphylococcus aureus which shows sequence homology to KlnA from Lactococcus lactis cremoris polypeptide. Histidine kinase is a component of the two component signal transduction systems (TCSTS), which is involved in bacterial pathogenesis and virulence. It undergoes autophosphorylation at a histidine residue and the phosphate group is transferred to the cognate response regulator. The polypeptide can be employed to produce antibodies and to identify agonists and antagonists which are used to prevent, inhibit or treat diseases, particularly Helicobacter pylori infections, such as gastric ulcers, gastrointestinal carcinoma, and gastritis. They

CC can also be used in vaccine formulations and for treating wounds and
 CC preventing infection by *S. aureus*. The histidine kinase products can be
 CC used to screen new antibacterial drugs effective against resistant
 CC *S. aureus* strains. The encoding polynucleotide may be used for staging of
 CC diseases, as reagents for screening genetic mutations and for genetic
 CC immunisation

XX Sequence 451 AA;

Query Match 100.0%; Score 2291; DB 2; Length 451;
 Best Local Similarity 100.0%; Pred. No. 8.7e-172;
 Matches 451; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTKRKLNNWIIVTMTFTFCLIIIFFLKDTLHNSLDDAERSDDINNLHFKSP 60
 DB 1 MTKRKLNNWIIVTMTFTFCLIIIFFLKDTLHNSLDDAERSDDINNLHFKSP 60
 QY 61 VKDISALDLNASLGNFQEIYDEHNKLFETSDNTVRVEPGYEHRYFDRVKKRYKGI 120
 DB 61 VKDISALDLNASLGNFQEIYDEHNKLFETSDNTVRVEPGYEHRYFDRVKKRYKGI 120
 QY 121 EYLIKEPITTDQPKGYSLLIHSLNVDNIVKSLYIIALAFGVIAITITATISYVFSTQI 180
 DB 121 EYLIKEPITTDQPKGYSLLIHSLNVDNIVKSLYIIALAFGVIAITITATISYVFSTQI 180
 QY 181 TKPLVSLSNKMIIRRDGFQNKQLQNTNVEEDNLTANTFNMMSQIEESFNQORQFVEDA 240
 DB 181 TKPLVSLSNKMIIRRDGFQNKQLQNTNVEEDNLTANTFNMMSQIEESFNQORQFVEDA 240
 QY 241 SHELTPLOIIOGHLNLIQWKKDPAVLSESNISIEENNRILKVEELLELTGKDND 300
 DB 241 SHELTPLOIIOGHLNLIQWKKDPAVLSESNISIEENNRILKVEELLELTGKDND 300
 QY 301 ISSEAQTVHINDEIRSRHSLKQLHPDYQPTDITLTSKNLEIKMKPHQFQOLFIFIDNAI 360
 DB 301 ISSEAQTVHINDEIRSRHSLKQLHPDYQPTDITLTSKNLEIKMKPHQFQOLFIFIDNAI 360
 QY 361 KYDVKNKKIKVTKLNKKKIETDHGIGIPEDQDFIDRFYRVVDKSRSGGNGLG 420
 DB 361 KYDVKNKKIKVTKLNKKKIETDHGIGIPEDQDFIDRFYRVVDKSRSGGNGLG 420
 QY 421 LSIAQKIQLNGSGSIKKSINKGTTFKIIIF 451
 DB 421 LSIAQKIQLNGSGSIKKSINKGTTFKIIIF 451

RESULT 2

ABU42408
 ID ABU42408 standard; protein; 451 AA.

XX AC ABU42408;

XX 19-JUN-2003 (first entry)

XX Protein encoded by prokaryotic essential gene #27935.

XX Antisense; prokaryotic essential gene; cell proliferation; drug design.

XX Staphylococcus aureus.

XX W0200277183-A2.

XX 03-OCT-2002.

XX 21-MAR-2002; 2002WO-US0009107.

XX 21-MAR-2001; 2001US-00815242.

PR 06-SEP-2001; 2001US-00948993.

PR 25-OCT-2001; 2001US-0342923P.

PR 08-FEB-2002; 2002US-00072851.

PR 06-MAR-2002; 2002US-0362699P.

XX (ELIT-) ELITRA PHARM INC.

XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX WPI; 2003-029926/02.
 DR N-PSDB; ACA46278.
 XX
 PT New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX
 PS Claim 25; SEQ ID NO 70332; 1766pp; English.

CC The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 451 AA;

Query Match 100.0%; Score 2291; DB 6; Length 451;
 Best Local Similarity 100.0%; Pred. No. 8.7e-172;
 Matches 451; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTKRKLNNWIIVTMTFTFCLIIIFFLKDTLHNSLDDAERSDDINNLHFKSP 60

DB 1 MTKRKLNNWIIVTMTFTFCLIIIFFLKDTLHNSLDDAERSDDINNLHFKSP 60

QY 61 VKDISALDLNASLGNFQEIYDEHNKLFETSDNTVRVEPGYEHRYFDRVKKRYKGI 120

DB 61 VKDISALDLNASLGNFQEIYDEHNKLFETSDNTVRVEPGYEHRYFDRVKKRYKGI 120

QY 121 EYLIKEPITTDQPKGYSLLIHSLNVDNIVKSLYIIALAFGVIAITITATISYVFSTQI 180

DB 121 EYLIKEPITTDQPKGYSLLIHSLNVDNIVKSLYIIALAFGVIAITITATISYVFSTQI 180

QY 181 TKPLVSLSNKMIIRRDGFQNKQLQNTNVEEDNLTANTFNMMSQIEESFNQORQFVEDA 240

DB 181 TKPLVSLSNKMIIRRDGFQNKQLQNTNVEEDNLTANTFNMMSQIEESFNQORQFVEDA 240

QY 241 SHELTPLOIIOGHLNLIQWKKDPAVLSESNISIEENNRILKVEELLELTGKDND 300

DB 241 SHELTPLOIIOGHLNLIQWKKDPAVLSESNISIEENNRILKVEELLELTGKDND 300

QY 301 ISSEAQTVHINDEIRSRHSLKQLHPDYQPTDITLTSKNLEIKMKPHQFQOLFIFIDNAI 360

DB 301 ISSEAQTVHINDEIRSRHSLKQLHPDYQPTDITLTSKNLEIKMKPHQFQOLFIFIDNAI 360

Db 301 ISSEAQTWHINDEIRSHLSKQLHPDYQFDLTLSKNLEIKMKPHQPEQLFLIPIDNAI 360
Qy 361 KYDVNKKIKVTRLNKKQKIIETDHGIGIPEEQDQDFIDRFYRVDSRSRSGGNGLG 420
Db 361 KYDVNKKIKVTRLNKKQKIIETDHGIGIPEEQDQDFIDRFYRVDSRSRSGGNGLG 420
Qy 421 LSIAQKIIQLNGGSIKIKSEINKGTTFKIIF 451
Db 421 LSIAQKIIQLNGGSIKIKSEINKGTTFKIIF 451

RESULT 3
ABM73032
ID ABM73032 standard; protein; 451 AA.
XX
AC ABM73032;
DT 20-NOV-2003 (first entry)
XX
DE Staphylococcus aureus protein #2272.
XX
KW Antibacterial; vaccine; gene therapy; infection; sepsis; diagnosis;
KW enzymatic assay; antibiotic target.
XX
OS Staphylococcus aureus.
XX
PN WO200294868-A2.
XX
PD 28-NOV-2002.
XX
PF 27-MAR-2002; 2002WO-IB002637.
XX
PR 27-MAR-2001; 2001GB-00007661.
XX
PA (CHIR-) CHIRON SPA.
XX
PI Masignani V, Mora M, Scarselli M;
XX
DR WPI; 2003-120786/11.
DR N-PSDB; ACF74592.
XX
PT New Staphylococcus aureus protein, useful as a vaccine for treating or
PT preventing Staphylococcal infection, specifically an infection caused by
PT S. aureus, e.g. sepsis.
XX
PS Claim 1; SEQ ID NO 4544; 49pp; English.
XX
CC The invention relates to novel genes and encoded proteins from
CC Staphylococcus aureus. A composition comprising the S. aureus protein, a
CC nucleic acid encoding the protein, or an antibody to the protein, is
CC useful as a pharmaceutical, particularly as a vaccine for treating or
CC preventing infection due to Staphylococcus bacteria, specifically an
CC infection caused by S. aureus. The composition is particularly useful for
CC treating or preventing sepsis in a patient. The composition can also be
CC used for diagnostics. The protein is also used in an assay for enzymatic
CC studies and as a target for antibiotics. This sequence represents one of
CC the novel S. aureus proteins of the invention
XX
SQ Sequence 451 AA;

Query Match 100.0%; Score 2291; DB 6; Length 451;
Best Local Similarity 100.0%; Pred. No. 8.7e-172;
Matches 451; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MTKKLRNNWIIVTMTFTVIFLCLIIIFLKDTHLSNLSLDAERSSDINNLFHSP 60
Db 1 MTKKLRNNWIIVTMTFTVIFLCLIIIFLKDTHLSNLSLDAERSSDINNLFHSP 60
Qy 61 VKDISALDNLASLGNFQBIIIYDEHNKLFETSNDNTVRPEGVEHYFDRVKKRYKGI 120
Db 61 VKDISALDNLASLGNFQBIIIYDEHNKLFETSNDNTVRPEGVEHYFDRVKKRYKGI 120
Qy 121 EYLIIKEPITTDQDFKGYSLIHSLENYDNIIVKSLYIIALAFGVIATITATISYVFSTQI 180

Db 121 EYLIIKEPITTDQDFKGYSLIHSLENYDNIIVKSLYIIALAFGVIATITATISYVFSTQI 180
Qy 181 TKPLVLSNKMIEIRRDGFQNKLOLNTNVEIDNLANTFNEMMSQIESEFNQORQFVEDA 240
Db 181 TKPLVLSNKMIEIRRDGFQNKLOLNTNVEIDNLANTFNEMMSQIESEFNQORQFVEDA 240
Qy 241 SHELRTPLOIITQGHNLNLIQRWGKDPVLEESLNIISIEEMNRIKLVLELLELTGQDVND 300
Db 241 SHELRTPLOIITQGHNLNLIQRWGKDPVLEESLNIISIEEMNRIKLVLELLELTGQDVND 300
Qy 301 ISSEAQTWHINDEIRSHLSKQLHPDYQFDLTLSKNLEIKMKPHQPEQLFLIPIDNAI 360
Db 301 ISSEAQTWHINDEIRSHLSKQLHPDYQFDLTLSKNLEIKMKPHQPEQLFLIPIDNAI 360
Qy 361 KYDVNKKIKVTRLNKKQKIIETDHGIGIPEEQDQDFIDRFYRVDSRSRSGGNGLG 420
Db 361 KYDVNKKIKVTRLNKKQKIIETDHGIGIPEEQDQDFIDRFYRVDSRSRSGGNGLG 420
Qy 421 LSIAQKIIQLNGGSIKIKSEINKGTTFKIIF 451
Db 421 LSIAQKIIQLNGGSIKIKSEINKGTTFKIIF 451

RESULT 4
ABP38305
ID ABP38305 standard; protein; 471 AA.
XX
AC ABP38305;
DT 24-JUL-2002 (first entry)
XX
DE Staphylococcus epidermidis ORF amino acid sequence SEQ ID NO:3150.
XX
KW Staphylococcus epidermidis; open reading frame; ORF; bacterial infection;
KW antibacterial; gene therapy.
XX
OS Staphylococcus epidermidis.
XX
PN US6380370-B1.
XX
PD 30-APR-2002.
XX
PF 13-AUG-1998; 98US-00134001.
XX
PR 14-AUG-1997; 97US-0055779P.
PR 08-NOV-1997; 97US-0084984P.
XX
(GENO-) GENOME THERAPEUTICS CORP.
XX
Doucette-Stamm LA, Bush D;
XX
WPI; 2002-381255/41.
DR N-PSDB; ABN90850.
XX
PT Novel isolated nucleic acid encoding a Staphylococcus epidermis
PT polypeptide, useful for diagnosing and treating bacterial infections.
XX
PS Disclosure; SEQ ID NO 3150; 267pp; English.
XX
CC ABN90538 to ABN93374 represent Staphylococcus epidermidis open reading
CC frame (ORF) nucleic acid sequences which encode the amino acid sequences
CC given in ABP35124 to ABP3960. The S. epidermidis sequences have
CC antibacterial activity and can be used in gene therapy. The sequences can
CC also be used in the diagnosis and treatment of bacterial infections,
CC particularly S. epidermidis infections. The sequences can be used to
CC screen for compounds able to interfere with the S. epidermidis life cycle
CC or inhibit S. epidermidis infection. N.B. The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from the USPTO web site
XX
SQ Sequence 471 AA;

```

Query Match      71.6%; Score 1640.5; DB 5; Length 471;
Best Local Similarity 69.9%; Pred. No. 1.5e-120;
Matches 316; Conservative 68; Mismatches 67; Indels 1; Gaps 1;

QY 1 MTKR-KLRNNWIVTMTITVFCLIIIFFLKDTLHNSLDDAERSSSDINNLFHSHK 59
DB 16 MIKRXKLKYYKWLITLTITFTILLFCLIIIFFLKDTLSRSEIDEAERSNDIANLFHSHK 75
QY 60 PVKDISALDNLASLGNFQEIILYVDHNNKULPETSNDNTVRVEPGYEHRYFDRVVKRYKG 119
DB 76 SLSDISALDNLASLENFQEIILYDDGRKLIQTSNDNTLAYDNKIDFKHPERIHQORSHG 135
QY 120 IEYLITKEPTTODFGYSLHLSLENYDNVKSLYIIALAFGVIATITATISYVFSTQ 179
DB 136 INYLVITEPIRSKDFSGYSLVLSLQNYDLNLSLYIIVALAFGLIATITAGVSYIFSSQ 195
QY 180 ITRPLVSLSNKMIERRDGFQNKQLQNTNYYEIDNLANTFNEMMSQIEESFNQORQFVED 239
DB 196 ITRPIVMSNMKNQIRRDGFQNKLELTNTNYYEIDNLANTFNEMMYQIEESFNQORQFVED 255
QY 240 ASHELRTPLQIIQGHNLNIQRWKKOPAVLEESNISIEEMNRIIKLVBELLBLTKGDVN 299
DB 256 ASHELRTPLQIIQGHNLNIQRWKKOPAVLEESNISIEEVRNITKLVBELLBLTKDRVN 315
QY 300 DISSEAQTVDHNDIERSRIHSLKQLHPDYQFDLTLSKNLEIKMKPHQFQLELIFIDNA 359
DB 316 HNVLECEENVVNSEIQSRVKSQHLHPDYTFETHLATKPIQLKINRHQFQLELLIFIDNA 375
QY 360 IKYDVNKKKIKVTRLNKKQKIIEITDHGIGIPEEDQDFIDRFYRVDKSRSRSQGGNGL 419
DB 376 MKYDTEHKHKIYVQLKNKMMIMDITDHGNGIPKADLEFIFDRFYRVDKSRSEARSQGGNGL 435
QY 420 GLSTAQKIQLNGSIKIKSEINKGTTFKIIF 451
DB 436 GLSTAEKIVOLNGMIQVESELOKYTTFKISF 467

RESULT 5
ID AAG82166 standard; protein; 284 AA.
AC AAG82166;
XX
XX
DT 03-SEP-2001 (first entry)
XX
DE S. epidermidis open reading frame protein sequence SEQ ID NO:1426.
KW Staphylococcus epidermidis SRI strain; infection; diagnosis; vaccination;
KW endocarditis.
XX
OS Staphylococcus epidermidis.
XX
PN WO200134809-A2.
XX
PD 17-MAY-2001.
XX
PF 09-NOV-2000; 2000WO-US030782.
XX
PR 09-NOV-1999; 99US-0164258P.
XX
XX (GLAX ) GLAXO GROUP LTD.
XX
XX Kimmerly WJ;
XX
XX WPI; 2001-316495/33.
XX
XX N-PSDB; AAH53016.
XX
PT Nucleic acids encoding polypeptides from Staphylococcus epidermidis,
PT useful for vaccinating against infections, e.g. endocarditis.
XX
XX Claim 18; Page 403; 2188pp; English.
XX
XX AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides

```

```

CC (II), given in AAG81454 to AAG83120, from Staphylococcus epidermidis. (I)
CC and (II) can have antibacterial activity and therefore can be used in
CC vaccination. The nucleic acids (I) may be used to produce the S.
CC epidermidis polypeptides (II) via the production of vectors containing
CC them which are used to produce hosts cells which express the
CC polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
CC used to vaccinate subjects and to raise antibodies against the bacteria.
CC The polypeptides may also be used to assay for other inhibitors of their
CC activity and therefore identify compounds that may be used for the
CC treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to
CC AAH55090 represent specifically claimed S. epidermidis genomic DNA
CC polynucleotide sequences from the present invention. AAH55091 to AAH55098
CC represent oligonucleotide sequences and primers which are used in the
CC exemplification of the present invention. N.B. The present invention
CC specifically claims all the polynucleotide sequences given in the
CC sequence listing of the present specification, however the sequence
CC listing only goes up to SEQ ID NO:4454 so even though sequences are given
CC in the disclosure for SEQ ID NO:4465 to 4472, no sequences are present
CC for SEQ ID NO:4455 to 4464
XX
SQ Sequence 284 AA;

Query Match      46.7%; Score 1069; DB 4; Length 284;
Best Local Similarity 73.2%; Pred. No. 7.6e-76;
Matches 205; Conservative 38; Mismatches 37; Indels 0; Gaps 0;

QY 172 ISYVFSTQITKPLVSLSNKMIERRDGFQNKQLQNTNYYEIDNLANTFNEMMSQIEESFN 231
DB 1 VSYFSSQITKPIVTMSNMKNQIRRDGFQNKLELTNTNYYEIDNLANTFNEMMYQIEESFN 60
QY 232 QORQFVEDASHELRTPLQIIQGHNLNIQRWKKOPAVLEESNISIEEMNRIIKLVEELL 291
DB 61 QORQFVEDASHELRTPLQIIQGHNLNIQRWKKOPAVLEESNISIEEVRNITKLVEELL 120
QY 292 ELTKGDVNDISSEAQTVDHNDIERSRIHSLKQLHPDYQFDLTLSKNLEIKMKPHQFQLE 351
DB 121 LLTKDRVNVHNVLECEENVVNSEIQSRVKSQHLHPDYTFETHLATKPIQLKINRHQFQLE 180
QY 352 FLIFIDNAIKYDVNKKIKVTRLNKKQKIIEITDHGIGIPEEDQDFIDRFYRVDKSR 411
DB 181 LLIFIDNAIKYDTEHKHKIYVQLKNKMMIMDITDHGNGIPKADLEFIFDRFYRVDKSRA 240
QY 412 RSQGGNGLGLSTAQKIQLNGSGSIKIKSEINKGTTFKIIF 451
DB 241 RSQGGNGLGLSTAEKIVOLNGMIQVESELOKYTTFKISF 280

RESULT 6
AAG82848
ID AAG82848 standard; protein; 284 AA.
AC AAG82848;
XX
XX
DT 03-SEP-2001 (first entry)
XX
DE S. epidermidis open reading frame protein sequence SEQ ID NO:2790.
XX
XX Staphylococcus epidermidis SRI strain; infection; diagnosis; vaccination;
KW endocarditis.
XX
OS Staphylococcus epidermidis.
XX
PN WO200134809-A2.
XX
PD 17-MAY-2001.
XX
PF 09-NOV-2000; 2000WO-US030782.
XX
PR 09-NOV-1999; 99US-0164258P.
XX
XX (GLAX ) GLAXO GROUP LTD.
XX
XX Kimmerly WJ;

```


Db 293 LDESINASLTERMKKLVQEMLDLSRAEQISQTKELQITDVNATVEQVRNFEVMEYENF 352
 Qy 329 QFDTDLTSKNLEIKWKPHQEPOLFLIFDINAIVKVNKKIKVTKRLKNKQKIIETDYG 388
 Db 353 TFLTKEDDTLRALIQNHLEQIIIMDNVAVKSGDGTVDHMYKEQKQIHDVDRYG 412
 Qy 389 IGPEEDQDFIDRFYRVDKRSRSGGNGGLGSLAQKIIQLNGSGSIKIKSEINKGTTFK 448
 Db 413 EGISQEBIDKIFNFRYRVDKRSRSGGNGGLGSLAQKQVGYLGTINAVSEPDKGTITK 472
 Qy 449 I 449
 Db 473 I 473

RESULT 8
 ABU33012
 ID ABU33012 standard; protein; 483 AA.
 AC ABU33012;
 XX
 DT 19-JUN-2003 (first entry)
 XX
 DE Protein encoded by Prokaryotic essential gene #18539.
 XX
 KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
 XX
 OS Listeria monocytogenes.
 XX
 PN MO200277183-A2.
 XX
 PD 03-OCT-2002.
 XX
 PF 21-MAR-2002; 2002WO-US009107.
 XX
 PR 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 XX
 PA (ELIT-) ELITRA PHARM INC.
 XX
 PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HJ;
 XX
 DR WPI; 2003-029926/02.
 DR N-PSDB; ACA36882.
 XX
 XX New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX
 XX Claim 25; SEQ ID NO 60936; 1766pp; English.
 XX
 CC The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene

CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 XX Sequence 483 AA;
 XX
 Query Match 30.7%; Score 703.5; DB 6; Length 483;
 Best Local Similarity 33.7%; Pred. No. 9.5e-47;
 Matches 162; Conservative 102; Mismatches 170; Indels 47; Gaps 10;
 Qy 1 MTKRKLNNWIIVTMTITFTFLFCLII-----TFFLKDTLHSE 41
 Db 8 LKRSRLKFKWTGASAAIFLTFELFSYAIYQGIGQMLNEEPEVKELLATTTLTNQD 67
 Qy 42 LDDAERSSSDINNLPHSKPVKDISALDNLASLGNFOEIIYDEHN--NKLPTSNDNTV 98
 Db 68 LTDNE---EIKYLFNNDKT-----VNRKLDQ-QVINLYDKDGHFINKYFYSRQDIT 115
 Qy 99 RVPEGYEHRYF----DRVI--KRYKGIIEYLIKEPIITQD---FKGYSLLIHSLNENYD 149
 Db 116 SID---FSQYFVSGTDFKFMNKPTIDGQKMTAQMEIVADDNTTVIGYQAVNPLTSYNR 172
 Qy 150 IVKSLYIIIALAFQVIATITATISYVFSTQITKPLVSLNKNMIEIRDFGQFNKLQNTN- 208
 Db 173 MMDRLVMTWLLGVALFISGMLGYLLAQNPLNPLRLARTMNDIRKGFQKRIETKTN 232
 Qy 209 YEBIDNLANTFNMMSQIEESFNQORQFVEDASHELRTPLQITQGHNLIQRWKKDPAV 268
 Db 233 RDEIGELTVVFNDMWTRITSEFQKQFVEDASHELRTPVQIMEGHLKLLTRWKKDPAV 292
 Qy 269 LEESLNISIEENRIIKLVEELELELTKGDVNDISSEAQTVDHINDEIRSRHSLKQLHPDY 328
 Db 293 LBSLNASLTELRLMKLVQEMLDLSRAEQISQTKELQITDVNATVEQVRNFEVMEYENF 352
 Qy 329 QPDTDLTSKNLEIKWKPHQEPOLFLIFDINAIVKVNKKIKVTKRLKNKQKIIETDYG 388
 Db 353 TFLTKEDDTLRALIQNHLEQIIIMDNVAVKSGDGTVDHMYKEQKQIHDVDRYG 412
 Qy 389 IGPEEDQDFIDRFYRVDKRSRSGGNGGLGSLAQKIIQLNGSGSIKIKSEINKGTTFK 448
 Db 413 EGISQEBIDKIFNFRYRVDKRSRSGGNGGLGSLAQKQVGYLGTINAVSEPDKGTITK 472
 Qy 449 I 449
 Db 473 I 473

RESULT 9
 ADC96804
 ID ADC96804 standard; protein; 521 AA.
 XX
 AC ADC96804;
 XX
 DT 01-JAN-2004 (first entry)
 XX
 DE E. faecium protein sequence SEQ ID 6431.
 XX
 KW Vaccine; urinary tract infection; bacteraemia; endocarditis; wound;
 KW abdominal-pelvic infection.
 XX
 OS Enterococcus faecium.
 XX
 PN US6583275-B1.

```
XX PD 24-JUN-2003.
XX PF 30-JUN-1998; 98US-00107532.
XX PR 02-JUL-1997; 97US-0051571P.
XX PR 14-MAY-1998; 98US-0085598P.
XX PA (GENO-) GENOME THERAPEUTICS CORP.
XX PI Doucette-Stamm LA, Bush D;
XX DR WPI; 2003-799836/75.
XX DR N-PSDB; ADC93150.
XX PT New isolated nucleic acid derived from Enterococcus faecium encoding an
XX PT Enterococcus faecium polypeptide useful for detection, prevention and
XX PT treatment of a pathological condition resulting from a bacterial
XX PT infection.
XX PS Example 1; SEQ ID NO 6431; 243pp; English.
XX PS The invention relates to an isolated nucleic acid derived from
XX CC Enterococcus faecium encoding an Enterococcus faecium polypeptide having
XX CC one of 10 fully defined sequences given in the (or comprising 40
XX CC sequential nucleotides chosen from any of the nucleic acids, its
XX CC complement or sequences hybridising to it). Also included are a
XX CC recombinant vector comprising the nucleic acid operably linked to
XX CC transcription regulatory element, a cell comprising the vector and a
XX CC single-stranded probe comprising the nucleic acid. The nucleic acids are
XX CC chosen from 3654 disclosed sequences encoding 3654 disclosed proteins.
XX CC The nucleic acids is useful for diagnosing pathological conditions
XX CC resulting from E. faecium bacterial infection (e.g. urinary tract
XX CC infection, bacteraemia, endocarditis, wounds and abdominal-pelvic
XX CC infection) and for screening drugs such as agonists and antagonists. The
XX CC nucleic acid is useful for recombinant production of Candida albicans -
XX CC derived peptides or antisense polypeptides. Pharmaceutical compositions
XX CC and vaccines containing the nucleic acid are useful for preventing or
XX CC treating Enterococcus faecium infections. The present sequence represents
XX CC one if the disclosed E. faecium proteins.
XX PS Sequence 521 AA;
XX PS Query Match 27.3%; Score 625; DB 7; Length 521;
XX PS Best Local Similarity 31.7%; Pred. No. 1.6e-40;
XX PS Matches 158; Conservative 108; Mismatches 165; Indels 68; Gaps 14;
Qy 3 KRKLNN-----NWIIIVTMTFTVIFLPCIIIPKDTLHNSLDDAERSSDI-----52
Db 29 KRELKGPSLTIKAFASFFIFVFTIPAVITYKSSINLIIVAKERNNVERTISEVASRLA 88
Qy 53 -----NNLFHSGPKVKDISALDNLASL--GNFQEI-----IIYDE 84
Db 89 NSDELTIVTNRNL--KATNDGDEVVDHTSLVSGNLMKIDSFAELGQPSMDLYVDL 146
Qy 85 HNNKLFETSDN-----TVRVEPGYEHRYFDRVI-----KKRYKGYEYLIIKEPITTDQFK 135
Db 147 DENLIFKTHENSRLIQTTRKVP-----TIVLECKTGFLSIQPIYSK---GTREKI 195
Qy 136 GYSLIIHLSLENDVIVKSLYIIAALAFGVIATITATISYVFSQTITKPLVSLSNKMEIR 195
Db 196 GYAQSFLSFSFYDIRNKLTLVLEVVSLILSGFLSSVPLKPLKVLRTDMTIR 255
Qy 196 RDGFQNL--OLANTVYEIDNLANTFNEMSQISEFNQORQFVEDASHELRPLQIIQ 252
Db 256 KDP-QSDIHMFEDTN--DELADLAEIFNEMJDRMLRYIEQQEQFVEDVSHLRTPEVAIE 313
Qy 253 GHLNLIORWGKDPVAVLEESLNISTEEMNRIIKLVLELLELTGKDNDNISSEAQTIVND 312
Db 314 GHLNLIORWGKDPVAVLEESLNISTEEMNRIIKLVLELLELTGKDNDNISSEAQTIVND 312
Qy 313 BIRSRHSLKQLHPDY--QFTDITLTSKNLEIKMKPHQEPQLFIDNAIKYDVKNKKIK 370
XX PS 374 VTYQVYNNFKILYPPYVITLDDLT-KEVTLQIYRNHFEEQIIIIILDNVAKYSTRKEWH 432
XX Qy 371 VKTRLKKNQKIIETDGHGIPEDQDFIFDRFYRVKDSRSRSGGNGLSIAQKIQL 430
XX Db 433 ISISSTLNEFEIAIQDFGEGIPKEDLNKIFNRFYRVKARATYKGGNGLGSLIAQLVEN 492
XX Qy 431 NGGSIKIKSEINKGTTFKI 449
XX Db 493 YKGRILAESVVGQGTIFRI 511
XX PS RESULT 10
XX PS ABB54939
XX ID ABB54939 standard; protein; 491 AA.
XX AC ABB54939;
XX XX 29-AUG-2003 (revised)
XX DT 16-MAY-2002 (first entry)
XX XX Lactococcus lactis protein kina.
XX DE Biosynthesis; biodegradation; lactic bacterium; yogurt; cheese.
XX KW Lactococcus lactis; IL1403.
XX OS FR2807446-A1.
XX PN 12-OCT-2001.
XX PD 11-APR-2000; 2000FR-00004630.
XX PF 11-APR-2000; 2000FR-00004630.
XX PR (INRG ) INRA INST NAT RECH AGRONOMIQUE.
XX PA Bolotine A, Sorokine A, Renault P, Ehrlich SD;
XX PI WPI; 2002-043418/06.
XX DR New nucleotide sequence useful in the identification of Lactococcus
XX PT lactis and related species.
XX PS Claim 6; SEQ ID NO 1641; 2504pp; French.
XX CC The present invention is related to a Lactococcus lactis nucleotide
XX CC sequence (AB90521) and related proteins (AB93300-AB93362). The nucleic
XX CC acid sequence is useful in the detection and/or amplification of nucleic
XX CC acid sequence, particularly to identify Lactococcus lactis or related
XX CC species. The proteins of the invention are useful for the biosynthesis or
XX CC biodegradation of a composition of interest. The invention helps research
XX CC in lactic bacteria, particularly useful in the production of yogurt and
XX CC cheese. Note: The sequence data for this patent is based on equivalent
XX CC patent WO200177334 (published 18-OCT-2001) which is available in
XX CC electronic format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences. (Updated on 29-AUG-2003 to
XX CC standardise OS field)
XX PS Sequence 491 AA;
XX PS Query Match 26.5%; Score 606.5; DB 5; Length 491;
XX PS Best Local Similarity 31.8%; Pred. No. 4.2e-39;
XX PS Matches 152; Conservative 100; Mismatches 187; Indels 39; Gaps 12;
Qy 3 KRKLNNWIIIVTMTFTVIFLPC-----LIIFPLKD--TLHNSELD-----DAE 46
Db 20 KRSIMLRWAFANVFCEITFTTFLATLYQLTISFKEEQQLLTRSMDSVSEVLEKADAP 79
Qy 47 RSSSDINNLHFSKPV---KDISALDNLASLGNFQEIILIIYDEHNNKLFETSDNTRVPEPG 103
Db 80 LNSNLSNLTATSKIQNGSESGMSLGSIIIGTRKAFYIYD-LNHKLLYSTNRHTF---G 134
Qy 104 YEHRYFDRVIRKRYKGYEYLIIKEPI--TTQDFKGYSLIIHLSLENDVIVKSLYIIAALAF 161
```



```

Db 122 DKQWIFTDNEES---SPGL-HGPIGRVYHDIHQVGRFSMTQKVKYSNRTGKFGVGVQV 177
Qy 141 IHSLENDVNIKSLYIIAALAFGVATITATISYVFSTQITKPLVLSNKMIEIRDDGQF 200
Db 178 FHDLGNYVIRARLLFWLLVVELFGTSLAYLILITTRRLKPLNLH-----EVMENISE 233
Qy 201 NKQLN-----TNVEETDNLANTFNEMMSOIESFNQORFVDSASHELTPLOIIOGHL 255
Db 234 NPNLNLRSIDSGDEIEELSVIFDNMLDKLETHKLQSFISDVSHELTPVAILNGHI 293
Qy 256 NLQRMGKKDPVAVLEESINISIEEMNRIIKLVELELTK-----GDVNDISSEAGTVH 309
Db 294 GLLRQMGKDSIDLEESLTATAHEADRMATMINDMLMIRVQGSFEGHQDM-----TV- 347
Qy 310 INDEIRSRHSLKQLHPDYQFDLTLSKNLEIKMKHPQFLIFIDNAIKYDVNKKI 369
Db 348 LEDSIETVGNFRVLREDFTFTWQSENPKTARIYKNHFEQALMILIDNAVKYSRKEKKI 407
Qy 370 KVTRLKNNKOK-IIEITDHGIGIPEEDODFIDRFYRVDKRSR--SQGGNGIGLSIAOK 426
Db 408 AINLSVTGQEAIVRQDKGEGSKEDIEHIFERFYRTDKSRNTSTQAGLIGLSILKQ 467
Qy 427 IIOLANGGSIKIKSEINKGTTT 447
Db 468 IVDGYHLQMKVESELNEGSVF 488

RESULT 14
ABP27834
ID ABP27834 standard; protein; 501 AA.
XX
AC ABP27834;
XX
DT 02-JUL-2002 (first entry)
XX
DE Streptococcus polypeptide SEQ ID NO 4844.
XX
KW Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
KW group A streptococcus; Streptococcus pyogenes; antibacterial;
KW antiinflammatory; infection; vaccine; meningitis; gene therapy.
XX
OS Streptococcus agalactiae.
XX
PN WQ20023471-A2.
XX
PD 02-MAY-2002.
XX
PF 29-OCT-2001; 2001WO-GB004789.
XX
PR 27-OCT-2000; 2000GB-00026333.
XX
PR 24-NOV-2000; 2000GB-00028727.
XX
PR 07-MAR-2001; 2001GB-00005640.
XX
PA (CHIR-) CHIRON SPA.
PA (GENO-) INST GENOMIC RES.
PI Telford J, Masignani V, Margarit Y Rosi, Grandi G, Fraser C;
PI Tettelin H;
XX
XX WPI; 2002-352536/38.
XX
XX N-PSDB; ABN68465.
XX
PT New Streptococcus protein for the treatment or prevention of infection or
PT disease caused by Streptococcus bacteria, such as meningitis, and for
PT detecting a compound that binds to the protein.
XX
XX Claim 1; Page 3646; 4525pp; English.
XX
XX The invention relates to a protein (ABP25413-ABP20895) from group B
XX streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
XX (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in
XX the specification. The proteins have antibacterial and antiinflammatory

```

```

CC activity. (I), nucleic acids encoding (I), ABN6044-ABN71526 and
CC antibodies that bind (I) are used in the manufacture of medicaments for
CC the treatment or prevention of infection or disease caused by
CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
CC Nucleic acids encoding (I) are used to detect Streptococcus in a
CC biological sample. (I) is used to determine whether a compound binds to
CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be
CC used as a vaccine or diagnostic composition. The disease caused by
CC Streptococcus that is prevented or treated may be meningitis. Nucleic
CC acid encoding (I) may be used to recombinantly produce (I) and may be
CC used in gene therapy. Antibodies to (I) are used for affinity
CC chromatography, immunoassays, and distinguishing/identifying
CC Streptococcus proteins
XX
XX Sequence 501 AA;
XX
Query Match 22.5%; Score 515; DB 5; Length 501;
Best Local Similarity 27.7%; Pred. No. 6.9e-32;
Matches 134; Conservative 106; Mismatches 169; Indels 74; Gaps 12;

Qy 20 VTIFFCLIIIFL-----KDTLHNSLD-----DAERSSSDINLPH 57
Db 23 VFILFSLFTVPFSLVYVTSATRYVLHREKINVGSRLEKTRVLSQANSLTSDILELY 82
Qy 58 SKPVKD-----ISALDLNASLGNFOBIIYDSHNNKLFETSNDNTVRVPGYE 105
Db 83 NQVFADDDIYPHKQNGIVTGESIDSILYVQEMTLYDVNRKPVF-----SLRTGMPTI 137
Qy 106 HRYDFRVIKKRYGIEYLIKEPITQ---DFKGYSLLIHSLNENYDNVKSLYIIAALFG 162
Db 138 GKSMGKVIISKVADMEGFGVGTAKIYSQKTGQLLVQVQIFVNLGRYYSMRQNIIVFLIMME 197
Qy 163 VIATITATISYVFSTQITKPLVLSNKMIEIRDDGQFNKLQNT---NYEERIDNLANTF 219
Db 198 VLGTVLALVINSATKRIVRPVKNLHDLHQISEN--PSNLEIRSKVRSEDEIGELSRIF 255
Qy 220 NEMMSOIESFNQORFVEDASHELTPLOIIOGHLNLIQRMGKKDPVAVLEESLNISIE 279
Db 256 DGMLDQLEDTYTRQSFISDVSHELTPVAVVKGHIGLQRMGKKDPEILERSLAAAYHE 315
Qy 280 MNRRIKLVEEL-----LELTGKDNDVNDISEAQTVHNDEIRSRHSLKQLHPDYQF 330
Db 316 ADRLMSLMINDMLNMRVQGSLELHQDEVDTLSSISW-----IENFRILREDDFQF 366
Qy 331 DTDLSKSLNLEI-----KMKHPQEQFLIFIDNAIKYDVNKKIKVTRLKNNKOKIIEITD 386
Db 367 ----IPENNISDIWVGKIYKIHFEQALMILIDNAIKYSPSYKESVSVLSDVNDNFATVVVKD 423
Qy 387 HGIGIPEEDODFIDRFYRVDKRSR--SQGGNGIGLSIAOKIIOGNGSIIKISEINKG 444
Db 424 KEGISDEDEIEFIDRFYRTDKSRNRESTQAGLIGLSVFKQIMDAYHLKVDIKSELNOG 483
Qy 445 TTF 447
Db 484 TEF 486

RESULT 15
ABP29835
ID ABP29835 standard; protein; 501 AA.
XX
AC ABP29835;
XX
DT 02-JUL-2002 (first entry)
XX
DE Streptococcus polypeptide SEQ ID NO 8846.
XX
KW Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
KW group A streptococcus; Streptococcus pyogenes; antibacterial;
KW antiinflammatory; infection; vaccine; meningitis; gene therapy.
XX
OS Streptococcus agalactiae.
XX

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 5, 2004, 07:05:52 ; Search time 1100.87 Seconds
(without alignments)
10135.683 Million cell-updates/sec

Title: US-09-006-627-1
Perfect score: 2201
Sequence: 1 TAAATTAAAGCACTATT.....AAAACAAATACAGTGGTT 2201

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 3340653 seqs, 2534783454 residues

Total number of hits satisfying chosen parameters: 6681306

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:
3: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:
13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:
14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:
15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:
16: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:
17: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:
18: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:
19: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1432	65.1	1440	8	US-08-781-986A-725
2	1432	65.1	1440	13	US-10-329-624-725
3	1356	61.6	1356	13	US-10-282-122A-34148
4	673.2	30.6	4858	8	US-08-781-986A-379
5	673.2	30.6	4858	13	US-10-329-624-379
6	399.8	18.2	495269	16	US-10-398-221-8
7	399.8	18.2	3011208	16	US-10-398-221-2058
8	246.6	11.2	1452	13	US-10-282-122A-24752
9	197.4	9.0	1351	9	US-09-070-927A-472
10	194.8	8.9	1308	9	US-09-070-927A-27
11	154.8	7.0	684707	16	US-10-398-221-9
12	154.8	7.0	3011208	16	US-10-398-221-2058
13	141	6.4	1449	13	US-10-282-122A-15787
14	136.8	6.2	1839	13	US-10-282-122A-9474

15	133.4	6.1	1755	13	US-10-282-122A-35146	Sequence 35146, A
16	131.8	6.0	690	13	US-10-282-122A-17336	Sequence 17336, A
17	130.2	5.9	760	13	US-08-781-986A-822	Sequence 822, App
18	130.2	5.9	760	13	US-10-329-624-822	Sequence 822, App
19	130	5.9	31517	9	US-09-070-927A-180	Sequence 180, App
20	128	5.8	1037	16	US-10-398-221-3163	Sequence 3163, App
21	128	5.8	1707	13	US-10-282-122A-16047	Sequence 16047, A
22	127.6	5.8	9425	8	US-08-781-986A-87	Sequence 87, Appl
23	127.6	5.8	9425	13	US-10-329-624-87	Sequence 87, Appl
24	126.8	5.8	1737	9	US-09-815-242-4265	Sequence 4265, App
25	126.8	5.8	1752	9	US-09-815-242-8499	Sequence 8499, App
26	126.4	5.7	717	16	US-10-398-221-886	Sequence 886, App
27	125.8	5.7	13022	9	US-09-070-927A-117	Sequence 117, App
28	125.6	5.7	699	9	US-09-925-637-45	Sequence 45, Appl
29	125.6	5.7	699	15	US-10-084-205-45	Sequence 45, Appl
30	125.6	5.7	699	17	US-10-712-713-45	Sequence 45, Appl
31	125.6	5.7	708	9	US-09-815-242-8902	Sequence 8902, App
32	125.6	5.7	708	13	US-10-282-122A-8093	Sequence 8093, App
33	125	5.7	721	13	US-10-282-122A-34637	Sequence 34637, A
34	124.6	5.7	705	13	US-10-282-122A-10198	Sequence 10198, A
35	124.4	5.7	696	13	US-10-282-122A-16259	Sequence 16259, A
36	124	5.6	699	9	US-09-815-242-4621	Sequence 4621, App
37	124	5.6	708	9	US-09-815-242-8576	Sequence 8576, App
38	123.6	5.6	1767	13	US-10-282-122A-8000	Sequence 8000, App
39	123.6	5.6	2244	9	US-09-866-232-1	Sequence 1, Appl
40	122.8	5.6	319630	16	US-10-398-221-7	Sequence 7, Appl
41	122.8	5.6	678	9	US-09-815-242-7381	Sequence 7381, App
42	122.2	5.6	678	13	US-10-335-977-2707	Sequence 2707, App
43	122.2	5.6	678	13	US-10-335-977-2708	Sequence 2708, App
44	120	5.5	663	16	US-10-398-221-732	Sequence 732, App
45	120	5.5	663	16	US-10-398-221-2755	Sequence 2755, App

ALIGNMENTS

RESULT 1

US-08-781-986A-725
; Sequence 725, Application US/08781986A
; Publication No. US20030054436A1
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5255
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781.986A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Benson, Bob
; REGISTRATION NUMBER: 30,446
; REFERENCE/DOCKET NUMBER: PB248PP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 725:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1440 base pairs

TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear

US-08-781-986A-725

Query Match 65.1%; Score 1432; DB 8; Length 1440;
Best Local Similarity 99.7%; Pred. No. 2.4e-241;
Matches 1432; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY	29	AAATCAAAATTTAAATATATACCTTATTAGAAAAGTCGTATATAGAGGTGACAAA	88
DB	4	AAATNCANTATTTAAATATATATACCTTATTAGAAAAGTCGTATATAGAGGTGACAAA	63
QY	89	TCACGCAAAATTTAATAGTAGAAGATGAACAAAACCTTAGCAAGATTTCTTGAATTTGGAAC	148
DB	64	TCACGCAAAATTTAATAGTAGAAGATGAACAAAACCTTAGCAAGATTTCTTGAATTTGGAAC	123
QY	149	TCACATGAAATTTACAAATGTGGACACAGAGATGATGGAACAAGCGGTTAGATAAAG	208
DB	124	TCACATGAAATTTACAAATGTGGACACAGAGATGATGGAACAAGCGGTTAGATAAAG	183
QY	209	CGCTTAGCCATTTACTATGATTTAATCATATTTAGATTTAATGTTGCGGTCAATTAATGCT	268
DB	184	CGCTTAGCCATTTACTATGATTTAATCATATTTAGATTTAATGTTGCGGTCAATTAATGCT	243
QY	269	TGAAATTTGTGCAAAAATTTAGACAAACAATCTACACCTATCATTAATAATACAGCGA	328
DB	244	TGAAATTTGTGCAAAAATTTAGACAAACAATCTACACCTATCATTAATAATACAGCGA	303
QY	329	AAAGTGATACGTATGACAAAGTTGCTGGCTTGATTAACGGTGACAGCATTAATAGTTA	388
DB	304	AAAGTGATACGTATGACAAAGTTGCTGGCTTGATTAACGGTGACAGCATTAATAGTTA	363
QY	389	AGCGTTTCATATTGAAAGACTTTTAGCAAGATTCGTGCAATTTTACGCTGACCCAC	448
DB	364	AGCGTTTCATATTGAAAGACTTTTAGCAAGATTCGTGCAATTTTACGCTGACCCAC	423
QY	449	AAAAGGATATTATCGATGTCAACGGTATTACAAATGATAAGAACGCTTTTAAAGTGACGG	508
DB	424	AAAAGGATATTATCGATGTCAACGGTATTACAAATGATAAGAACGCTTTTAAAGTGACGG	483
QY	509	TAAATCGCGCAGAAATTTGAATTAACAAAACAGAGATGATGATTAATCTCTAGCTG	568
DB	484	TAAATCGCGCAGAAATTTGAATTAACAAAACAGAGATGATGATTAATCTCTAGCTG	543
QY	569	AAAATAAAACCATGTTATGCAACGGGAACAAATTTTAAATCATGTATGGGGTTAATA	628
DB	544	AAAATAAAACCATGTTATGCAACGGGAACAAATTTTAAATCATGTATGGGGTTAATA	603
QY	629	GTGAAGTAGAAACAAATGTCGTAGATGTTTATATAAGATATTACGAAACAAAGTTAAAC	688
DB	604	GTGAAGTAGAAACAAATGTCGTAGATGTTTATATAAGATATTACGAAACAAAGTTAAAC	663
QY	689	CATACGATCGTGACAAAATGATTTGAAACAGTTCGTCGGCTGGGTATGTAGATGAC	748
DB	664	CATACGATCGTGACAAAATGATTTGAAACAGTTCGTCGGCTGGGTATGTAGATGAC	723
QY	749	AAAACGTAATTCGCGCAATTAACGATTTATGTTACACGATGATTAACGTTTGTACCGAT	808
DB	724	AAAACGTAATTCGCGCAATTAACGATTTATGTTACACGATGATTAACGTTTGTACCGAT	783
QY	809	ATTTTGTGTTTGTAAATTTATTTTCTTGAAGATACACTGCATAATAGTGAGCT	868
DB	784	ATTTTGTGTTTGTAAATTTATTTTCTTGAAGATACACTGCATAATAGTGAGCT	843
QY	869	TGATGATGCAACGAAGCTCAACGATATTAATAATTTTATTTCTTAAGCTGTTAA	928
DB	844	TGATGATGCAACGAAGCTCAACGATATTAATAATTTTATTTCTTAAGCTGTTAA	903
QY	929	AGATATCTGCATTTAGACTTGAATGATCTTTAGGTAATTTTCAAGAGATTAATTTTA	988
DB	904	AGATATCTGCATTTAGACTTGAATGATCTTTAGGTAATTTTCAAGAGATTAATTTTA	963

QY	989	TGATGAGCATATAATAAATTTATTTGAGACATCGAATGATAACACAGTGAGAGTTGAACC	1048
DB	964	TGATGAGCATATAATAAATTTATTTGAGACATCGAATGATAACACAGTGAGAGTTGAACC	1023
QY	1049	AGTTTATGAACACCGTTATTTTGACCGCGTAAATAAATAAAGCGCTATAAAGCGATTGAATA	1108
DB	1024	AGTTTATGAACACCGTTATTTTGACCGCGTAAATAAATAAAGCGCTATAAAGCGATTGAATA	1083
QY	1109	TTTAATTTAAGAAACCAATTTACAAACGCAAGATTTTCAAAAGGGGTATAGCTTTGTTAATTC	1168
DB	1084	TTTAATTTAAGAAACCAATTTACAAACGCAAGATTTTCAAAAGGGGTATAGCTTTGTTAATTC	1143
QY	1169	TTTACTAGAAAATTTATGATAACATCGTAAATAATTTGATATATCATTTGGCTGGCATTGG	1228
DB	1144	TTTACTAGAAAATTTATGATAACATCGTAAATAATTTGATATATCATTTGGCTGGCATTGG	1203
QY	1229	AGTGATTCACAAATTTAACTGCGCAACATCAGTTATGATTTTCAACACAAAATTTACTAA	1288
DB	1204	AGTGATTCACAAATTTAACTGCGCAACATCAGTTATGATTTTCAACACAAAATTTACTAA	1263
QY	1289	ACCGCTTGTGAGTTTATCAATAAATGATTTGAGATTCGACGAGATGGTTTTCAAAATAA	1348
DB	1264	ACCGCTTGTGAGTTTATCAATAAATGATTTGAGATTCGACGAGATGGTTTTCAAAATAA	1323
QY	1349	ATTCGCAATTTAATAAATTTATGAAGAAATAGATAAATTTAGCAAAATACGTTTAATGAGAT	1408
DB	1324	ATTCGCAATTTAATAAATTTATGAAGAAATAGATAAATTTAGCAAAATACGTTTAATGAGAT	1383
QY	1409	GATGAGCAAAATTTGAAGATCAATTTTATCAACAAAGCAATTTTGTGAAGATGGCTC	1465
DB	1384	GATGAGCAAAATTTGAAGATCAATTTTATCAACAAAGCAATTTTGTGAAGATGGCTC	1440

RESULT 2

US-10-329-624-725
Sequence 725, Application US/10329624
Publication No. US20040043037H1

GENERAL INFORMATION:

APPLICANT: Charles Kurych
Gil H. Cho
Patrick S. Dillon
Craig A. Rosen
Steven O. Barash
Michael R. Fannon
TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
NUMBER OF SEQUENCES: 5256
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/329,624
FILING DATE: 27-Dec-2002

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/956,171
FILING DATE: October 20, 1997
APPLICATION NUMBER: 60/009,861
FILING DATE: January 5, 1996
APPLICATION NUMBER: 08/781,986
FILING DATE: January 3, 1997

ATTORNEY/AGENT INFORMATION:

NAME: Mark J. Hyman
REGISTRATION NUMBER: 46,789
REFERENCE/DOCKET NUMBER: PB248P1D1
TELECOMMUNICATION INFORMATION:

